Milk production analysis with a flexible frontier production model

Tarmo Niemi
Master's thesis
Department of Economics
and Management
University of Helsinki
May 2020

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildingsprogram – Degree Programme
Faculty of Agriculture and Forestry	Agriculture economics

Tekijä – Författare – Author

Tarmo Niemi

Työn nimi – Arbetets titel – Title

Milk production analysis with a flexible frontier production model

Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track

Agriculture economics

Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's thesis	May 2020	65

Tiivistelmä – Referat – Abstract

The purpose of the study was to find a model to examine the effectiveness of feeding through monitoring calculations (MoC), while also investigating the total protein and energy responses of the feed diet. The mutual effectiveness of the environmental factors was also examined.

Today, as a result of changes in pricing, the milk protein and fat content are increasingly affecting the dairy's financial performance. The volatility of price for milk producer calculated based on the contents of the MoC used in the statistical analysis was 39,2 per cent if the pricing model 2018 of Cooperative Tuottajain Maito is used. However, nowadays the efficiency of milk production is mainly evaluated through the amount of milk. This study seeks to determine what the added value a milk yield efficiency assessment could bring to dairy farms.

The exceptionally dry year of 2018 influenced the comparison of inefficiencies between production inputs, for which statistics have been collected. From silage analyses, it can be concluded that, due to drought, farms have aimed at maximizing yields at the expense of the D value. This is particularly evident in the results of clover silage analyses. The aim Was to improve the comparability of results by considering the effect of the D value.

The material used in the thesis is a sample of the 2018 ProAgria milk production MoC, compiled from Valio farms. The statistics to be examined consist of 1 948 MoC: s to which 2 829 silage analyses can be attached. The analyses must be used to answer the research questions. MoC:s includes 91 984 individual observations of cows.

Testing of the Stochastic Production Frontier Function started from the Cobb-Douglas function, ending in a translog model based on the likehood ratio test. The model obtained is best suited to describe the statistics to be analysed. The model used in the thesis makes it relatively easy to set up the MoC:s, or alternatively, the dairy farms in order of efficiency. Most of the input data is available from the MoC:s and supplementary data from the regional cooperatives. But the fitting variable used to calculate is not enough to describe the variation in input prices in the analysis.

Several environmental factors were used to test the effect of feeding, barn, milking type and sharing of concentrated feed on milk yield. The differences between the different options were not so large that their output impact outweighed the expected cost of investments. It is worth looking at the impact on returns when the investment is timely for other reasons. At the same time, the effect of the silage harvest and the clover content of the silage on the milk yield was evaluated. The clover silage had a slightly positive impact if the effect of the D value was considered.

One contribution of this thesis is the emphasis on technical efficiency in the evaluation of milk production performance. The current method of assessing the efficiency of milk production is largely based on the measurement of milk volumes. Efficiency is thus based on the annual milk amount of the cows or the annual energy-corrected milk amount. However, through quality pricing, milk levels have an increasingly significant impact on milk yield. Environmental factors affect the analysis results when testing other environmental factors as well as the actual analysis model. Their interaction must therefore be measured.

Finally, the most significant contribution of the thesis is to apply protein and energy as inputs instead of different forages. And the following economic assessments. The response of energy was determined 0,94 and the one of protein was determined 0,19. The mean of $MRTS_{e,p}$ was -23,0 and the cost minimizing price ratio was around 115.

Avainsanat – Nyckelord – Keywords

Stochastic production frontier function, dairy farm, monitoring calculation, efficiency analysis, environmental factor.

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors

Professor Timo Sipiläinen

Säilytyspaikka – Förvaringställe – Where deposited

Tiivistelmä

Tutkimuksen tarkoituksena oli löytää malli ruokinnan tehokkuuden tutkimiseksi seurantalaskelmien avulla samalla kun tutkittiin myös rehun proteiini- ja energiakomponenttien tuotosvasteita. Myös olosuhdemuuttujien vaikutusta tutkittiin.

Nykyään hinnoittelumuutosten seurauksena maidon proteiini- ja rasvapitoisuus vaikuttaa yhä enemmän maitotilan taloudelliseen tulokseen. Tilastollisessa analyysissä käytetyn seurantalaskelmien maidon pitoisuuksien perusteella laskettu maidontuottajan hinnan volatiliteetti oli 39,2 prosenttia, jos käytetään Osuuskunnan Tuottajain Maidon hinnoittelumallia 2018. Nykyään maidontuotannon tehokkuutta arvioidaan kuitenkin pääasiassa maidon määrän perusteella. Tämän tutkimuksen tarkoituksena oli selvittää, mitä lisäarvoa maidontuoton tehokkuuden arviointi voisi tuoda maitotiloille.

Poikkeuksellisen kuiva vuosi 2018 ja toisaalta poikkeuksellisen sateinen vuosi 2017 vaikutti tehokkuusvertailuun tuotantopanosten välillä. Säilörehun analyysien perusteella voidaan päätellä, että kuivuuden vuoksi tilat ovat pyrkineet maksimoimaan sadon määrää D-arvon kustannuksella. Tämä näkyy erityisen selvästi apilarehun analyysien tuloksissa. Tavoitteena oli parantaa tulosten vertailtavuutta ottamalla huomioon D-arvon vaikutus.

Opinnäytetyössä käytetty materiaali on ProAgrian seurantalaskelmista valiolaisilta tiloilta koottu tilasto vuodelta 2018. Tutkittava tilasto koostui 1 948 seurantalaskelmasta, joihin voitiin liittää 2 829 säilörehuanalyysia. Analyysejä on käytetty vastaamaan tutkimuskysymyksiin. Seurantalaskelmat sisälsivät 91 984 yksittäistä lehmähavaintoa.

Stokastisen tuotantomenetelmän testaaminen alkoi Cobb-Douglas-funktiosta, päättyen translog-malliin. Saatu malli sopii parhaiten analysoitavien tilastojen kuvaamiseen. Opinnäytetyössä käytetyn mallin ansiosta seurantalaskelmien tai vaihtoehtoisesti tilojen tehokkuusvertailu oli suhteellisen helppoa. Suurin osa lähtötiedoista oli saatavana seurantalaskelmista ja hinnoittelutiedot alueosuuskunnilta. Laskemiseen käytetty sovitusmuuttuja ei riittävästi kuvannut panoshintojen vaihtelua analyysissä.

Useita olosuhdemuuttujia käytettiin kuvaamaan ruokinta-, navetta- ja lypsytyypin ja väkirehun jakotavan vaikutusta maitotuottoon. Eri vaihtoehtojen väliset erot eivät olleet niin suuria, että niiden tuotosvaikutukset olisivat suuremmat kuin investointien odotetut kustannukset. Tarkastelu kannattaa siinä vaiheessa, kun investointi on ajankohtainen muista syistä. Samanaikaisesti arvioitiin säilörehun korjuukerran ja apilapitoisuuden vaikutusta maidontuotantoon. Apilarehulla oli lievästi positiivinen vaikutus, jos D-arvon vaikutus otetaan huomioon.

Yksi tämän tutkielman kontribuutio on painotus tekniseen tehokkuuteen maidontuotannon suorituskyvyn arvioinnissa. Nykyinen maidontuotannon tehokkuuden arviointimenetelmä perustuu suurelta osin maidon määrien mittaamiseen. Tehokkuus perustuu siis lehmien vuotuiseen maitomäärään tai vuotuiseen energiakorjattuun maitomäärään. Maitotuottoon on kuitenkin laatuhinnoittelulla yhä merkittävämpi vaikutus. Olosuhdemuuttujilla on yhteisvaikutuksia analyysituloksiin, ja eri olosuhdemuuttujien vaikutukset voivat olla myös toisiinsa nähden päinvastaisia. Siksi niiden vuorovaikutusta on analysoitava.

Lopuksi, tutkielman merkittävin panos on proteiinin ja energian käyttäminen tuotantopanoksina erilaisten rehujen sijasta. Ja näiden panosten vaikutusten ekonometrinen mittaaminen mahdollisimman joustavan mallin avulla. Energian responssi oli 0,94 ja proteiinin responssi 0,19. Näiden tekninen rajakorvaavuussuhde ($MRTS_{e,p}$) oli -22,98 ja kustannuksia minimoiva hintasuhde noin 115.

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Table of Contents

1 Introduction

1.1 Background of the thesis

In Finland, the structural development of agriculture has continued since joining the European Union. This is partly due to the benefits of economies of scale and, in part, the retirement of the older generations. Those factors affect the continuity of the farm, especially at the point of transition when the change of generation becomes relevant. Often, the investments that would be necessary to secure the viability of a farm are not feasible on small farms based on economic realities. In these cases, the farm shall cease production, and fields are sold or leased. In dairy production, structural changes have led to increasing farm sizes with more cattle in the 2000s. Only in the size classes of more than 74 cows has the absolute number of cows and farms increased. The chart below illustrates structural developments from 2000 to 2018.

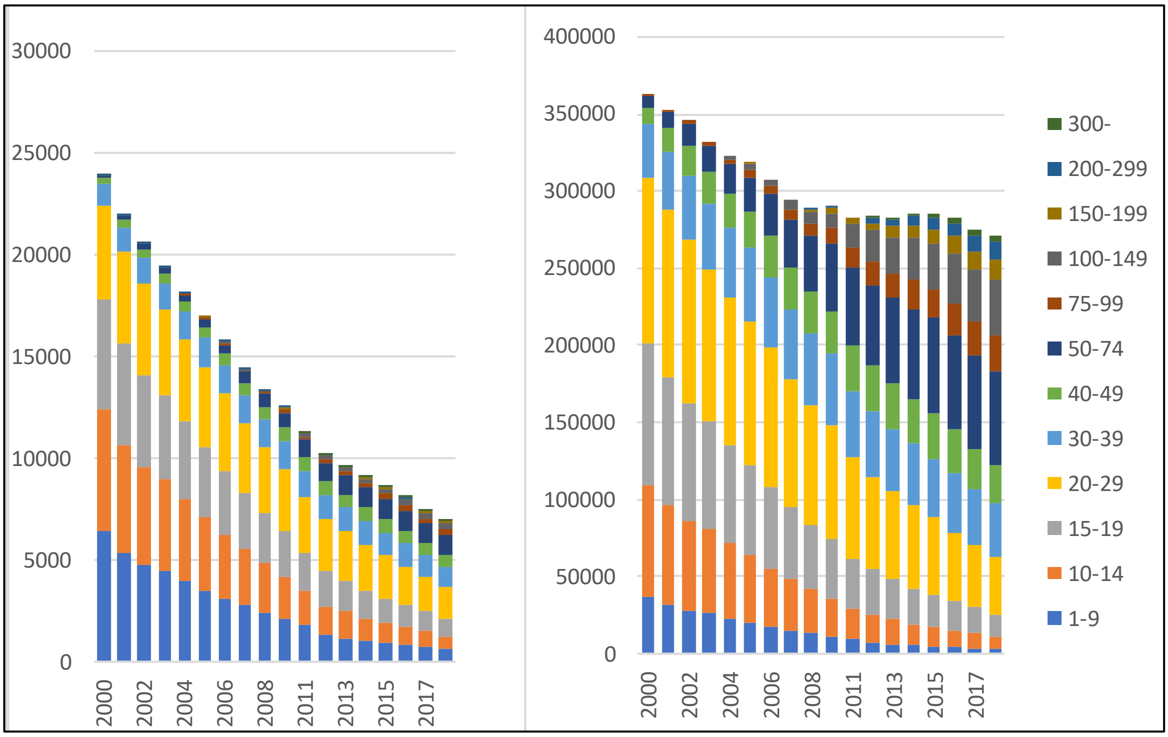


Figure 1. Distribution of farms (left panel) and cows (right panel) by size class 2000 - 2018 (Luke 2020).

Figure 1 illustrates the displacement of both farms and cows to the largest size ranges. The graph on the left demonstrates the evolution of the number of dairy farms among the different size classes. The graph on the right, in turn, shows the distribution of the number of cows among the different size classes. The colour codes of size are explained on the right. The growing size classes, with more than 74 cows, accounted for 7 % of the farms, totalling 435 in 2018. However, these farms represented 32,6 % of the total number of cows. At the end of 2019, there were 5 783 dairy farms, compared with 21 564 at the end of 2000 and 44 469 at the end of 1990. (Luke 2020.)

The development described above has led to changes in the management of cows, and at the same time has posed new challenges for assessing the economic performance of farms and thus foreseeing their activities. The focus has shifted from feeding and care measures for individual cows towards groups, which requires the ability to monitor milk production at the level of the herd. However, technical progress has made it also possible to follow singular cows as part of a group. Current milking systems provide enough up-to-date cow-specific information to make decisions for single cows. The increased impact of milk protein and fat content on milk returns underlines the importance of controlling production.

This Master’s thesis examines the current state of milk production and the challenges arising from the changes described above. This study assesses technical efficiency of milk production at the level of the whole herd. The average milk yield of cows is then compared with the average input used, which in this case is the amount of crude protein and energy in the feed. Correspondingly, milk return is calculated separately for each monitoring calculation (MoC). This milk return consists of the quantity of milk and the producer price of milk, which are affected by the protein and fat contents of the milk through coefficients. Quantitative research methods were used, and the writing follows the guidelines by Hirsjärvi, Remes and Sajavaara (1997).

1.2 Aim of the thesis and the research questions of the research

The purpose of this study was to find a model that can be used to examine the variation of technical efficiency of feeding through MoC:s. At the same time, the aim was to specify a fitting value that can estimate the effect of feed protein content variation on productivity. There are no input prices for purchased feed and no forage production costs, but input prices are crucial to the farm’s economic performance. The fitting value presented in the thesis contributes to illustrating the effect of silage protein content on efficiency.

The year when the impact of inputs on production efficiency is exceptionally dry 2018, for which the statistics have been collected (Finnish Meteorological Institute (FMI) 2018). From silage feed analysis, it can be concluded that due to drought, farms aimed at maximizing the yield of forage at the expense of D-value. This is particularly evident in the results of clover silage analyses. The issue was to improve the comparability of results by normalizing the results based on the D-value.

However, one year does not allow far-reaching conclusions on the differences in the effectiveness of the inputs.

Today, as a result of changes in pricing, the milk protein and fat content are increasingly affecting the dairy farm’s financial performance. The milk producer price volatility calculated based on the protein and fat content of milk was 39,2 %, if the milk pricing model of the Cooperative Tuottajain Maito is used. However, nowadays the productivity and efficiency of milk production is mainly evaluated through the amount of milk. This study seeks to determine what added value a milk return efficiency assessment could bring to dairy farms.

Relatively small numbers of animals are typically used in studies to determine the effect of a single variable. The significance of the individual observations is then emphasized. Correspondingly, the research environment seeks to exclude all distractions. In this paper, external distractions are not outlined, but the extent of the material contributes to preventing individual observations from affecting the outcome. This thesis aimed to determine whether the result of the studies and the results of the analysis based on the MoC:s are consistent.

The study will answer the following questions:

1. What are the responses to feed protein and energy and what are the differences in productivity and efficiency between MoC:s?
2. What environmental factors affect the differences in productivity and efficiency?
3. How can the interdependence of environmental factors influence the conclusions?

2 Previous studies and concepts

2.1 The concept of productivity based on monitoring calculations

Traditionally, the efficiency of milk production has been examined at the level of individual cows. The yield of each cow has been measured at monthly intervals and the milk concentrations per cow at least every two months. These measurements have determined both the average yield of dairy cattle and the annual yield per cow. The calculations monitored twice a year have been used to estimate the daily feed consumption of cows. Due to the comparability of the annual milk yield of cows, the fat content of the milk standardizes at 4 %. This means that a cow with a milk fat content of more than four per cent has energy corrected milk (ECM) amount higher than a cow with a milk fat content of less than 4 %. The same calculation model is used for the comparison of the average herd yield.

ProAgria defines the productivity of dairy farms through the farm condition model. In this model, it is possible to classify the holding according to, for example, the amount of milk sold to the dairy or the type of milking system. After the selection made, the model places the farm based on the different characteristic of the result into a group formed according to the resultant comparison farms. For example, based on milk production, the farms are grouped by quantity into the best 10, 30, 50, 70 or 90 %. Other production-based classification features are the percentage of milk sold to dairy, milk protein and fat percentage, number of cells[1], average milk yield and average first year's milk yield. Additionally, there are classification features like the mean number of cows, heifers' calving age[2], calving interval[3], inseminations per calving[4] and total hereditary value[5]. The economical classification features are the cost of producing milk, the milk return reduced by the cost of feed and the cost of purchased fodder. These are calculated per litre sold to the dairy. Well-being features include average number of calving per cow[6], mean amount of milk produced over lifetime of cows,

---

[1] Cells: The value measured on a sample of cow's milk from which the health status of the cow can be inferred. The number of cells measured in tank milk influences milk quality pricing (E, 1 and 2 grades).
[2] Heifers' calving age: Heifers normally breed at about two years of age. Increasing the calving age increases costs.
[3] Calving interval: Studies have shown that cows would be profitable for parturition every year. Increasing the calving interval weakens profitability.
[4] Inseminations per calving: Cows should be pregnant with as few inseminations as possible. The best 10% survive on average with about 1.5 inseminations per calving.
[5] Total hereditary value: Total hereditary value is the average value of the herd. The hereditary value of each cow is determined based on hereditary characteristics and its own characteristics. Based on this value, the animal material of the farm can be developed
[6] Average number of parturitions per cow: The value is calculated as the average number of parturitions per dairy cows on the farm. It tells about the durability of cows

longevity of cows, cow removal rate[7], and calf mortality[8]. Separate statistics are compiled for each characteristic, so the farm can be in the top 90 % based on the milk sold to the dairy and at the same time top 10 % based on the milk protein content.

2.2 Measuring of the productivity of cows

In Finland, monitoring milk yield is based on the International Committee for Animal Recording (ICAR) bylaws (ICAR 2019). The calculation methods for ProAgria's dairy herd bylaws are the milk amount of 305 days, the total milk amount, the milk amount of 12 months and annual milk amount (ProAgria 2019). The milk amount based on the 305 days per cow is the amount of milk produced for 305 days after calving. The total milk yield is the total amount of milk produced between the birth and the removal, regardless of the time, and the 12-month yield includes 12 consecutive calendar months, regardless of the year. Similarly, the annual yield is calculated as the quantity of milk produced in one calendar year and the average 305 days of livestock yields the mean amount of milk produced in 305 days. The average yield of the first-year cows consists of the milk produced by first-year cows during the calendar year in proportion to the number of days of production.

The practical measurement of milk volume is the responsibility of the livestock owner and depends on the type of milking system. When milking in a bucket, the milk is poured into a measuring bucket, whereupon a sample of milk is taken before being poured into farm tank. When milking with a pipeline milking system, a separate measuring pod is installed in the milking unit during which a small standard part of the milk to be milked is direct. The scale in the measuring pod shows the amount of milk. If necessary, a sample of milk is taken from the measuring pod as it represents the whole milk. The fat and protein concentration change during milking. These measuring methods are used in byres as a result of the milking system whereby the cattleman milks cows at the tied-housing. In a bucket and pipeline milking system, the measurement is based on the volume of milk, and the yield estimated by weight thus the volume must be converted by a factor into a weight.

In the loose-housing, the cows move freely within a restricted area, where they are milked in separate area at the milking parlor or by an automatic milking system (AMS) located in the same area as the cows. These milking systems generally include continuous measurement of the amount of milk bases on weight. For example, the DeLaval milking parlor has a fixed measuring device, which

---

[7] Cow removal rate: The depreciation percentage is calculated as the percentage of cows removed from the total number of cows during the year.
[8] Calf mortality: Percentage of total number of dead calves over one year.

considers milk by 150 g quantities of milk from a cow. The device weighs the remaining amount of milk separately. Each manufacturer has a slightly different measuring system. Usually, milking systems are connected to either a processor or a computer that records milking data. These processors or computers guide the feeding on the basis on milking data.

2.3 Economic optimization of the feeding

There is a great deal of research available on the economic impact of dry period feeding of dairy cows, and the results are convergent. High concentrated feeding level at dry period ratio is not economically viable. A study conducted by Agricultural University of Sweden used 24 cows representing the Swedish red cow breed. According to this study, high dry matter feeding during the dry season can negatively affect the overall milk yield, because high feed costs during dry season are not necessarily related to high milk yield after calving (Agenäs 2003). Only fat content responds positively to high feeding during the dry season (Agenäs 2003). According to a feeding experiment conducted by the Department of Animal Science at University of Helsinki in 2002, the share of concentrated feed in the diet before calving had no effect on the post-calving silage dry matter intake (Kokkonen, Tesfa, Tuori & Syrjälä-Qvist 2004). Thirty Friesian multiparous cows were divided into six groups. During the first five weeks after calving, a rapid increase in concentrated feed increased milk yield but decreased energy efficiency and decreased milk fat content (Kokkonen et al. 2004).

A study in Hamilton, New Zealand, found that low ratio of concentrated feed at dry period reduced fat and protein content in milk and energy-corrected milk amount (ECM) during the first five weeks after calving, but did not affect the amount of milk for the whole lactation period (Roche 2007). This research consisted of 64 multiparturition cows and was conducted by the Ruakura Animal Ethics Committee. In this paper, the feeding of cows based on grazing and the level of feeding was adjusted by pasture rotation[9]. Simultaneously, it was found that high dry period ratio feeding exposes cows with paralysis[10] and ketosis[11]. Low dry period ratio can be offset by high nutrient intake during the early lactation period[12] with a high energy and digestibility (Dewhurst et al. 2000; Agenäs 2003).

---

[9] Pasture rotation: The pasture rotation regulates the feeding of cows by transferring the cows to the new pasture according to their feeding needs. Used especially in New Zealand and Ireland, where milk production is based on grazing.
[10] Paralysis: Paralysis is a condition of calcium deficiency at calving which, when left untreated, leads to cow immobility and death.
[11] Ketosis: Ketosis is an energy deficiency in cows, causes weight loss and reduces milk production
[12] Lactation period: During the lactation period, the cow produces milk.

According to the Institute of Grassland and Environmental Research, Aberystwyth (UK), obesity due to high dry period ratio feeding can be detrimental to cow fertility (Dawhurst et al. 2000). The subject was 48 Holstein Friesian cows. Vanholder (2015) highlighted the risk of ketosis during the second week after calving increased with the prolongation of previous lactation period, especially in the case of older cows that produced a lot of colostrum[13] immediately after calving. The study of Vanholder (2015) carried out in Vorden, the Nederlands and included 1 715 cows in 23 dairy farms.

Restrictions on feeding during the lactation period significantly reduce milk production throughout the restrictive period and ten weeks thereafter (Roche 2007). A study by Agroscope Liebefeld-Posieux Research Center in Switzerland showed that the amount of colostrum produced after calving does not predict the amount of milk produced throughout the production season (Kessler 2014). The investigation included 56 Holstein cows, of which 17 were calving first time and 39 more than one time. The average parturition ratio of the whole group was 2,9 and the average milk yield was 8 137 kg. The cows were fed according to their energy and nutrient requirements.

The effect of fiber content, digestibility and partial size of silage on the feeding time variation of the cow is more than one hour a day (Grant & Ferraretto 2018). Their paper pointed that especially high-yielding cows may become the intake of their nutrient endangered if their dry matter intake is limited by prolonged silage chewing time when eating (McLeod, Kennedy & Minson 1990). At the same time, the resting and ruminating time is reduced as more time is spend at the feed bunk (Grant & Ferraretto 2018). A summary of feeding studies of legumes, grasses and straw by Mertens (1997) compared the time taken to chew the feed of dairy cows and other ruminants. He discovered that the time taken to chew the same size fiber particles ranged from 111 minutes to 209 minutes per kilogram.

In economic terms, as a summary of previous analysis, it is conceivable that abundant feeding in the dry period and prolonging that period will increase costs that cannot be compensated for during the lactation period. At the same time, the risk of loss of production due to high dry period ratio feeding is increased through parturition defects, ketosis and fertility problems. Restrictions on feeding during the lactation period have a greater impact on total milk amount than during the dry period, and the effect continues after restriction as reduced milk yield.

---

[13] Colostrum: Colostrum is a high-content-fat milk produced by the cow for approximately five days after calving.

2.4 Assignment of the responses of forages

In the current ruminant feeding system, as last amended in 2010, the energy value expresses megajoules of convertible energy according to the SI system. The change was made to facilitate the international comparability of feed values, but system differences between countries prevent direct comparison of feed values. Only about a quarter of the energy bound to feed can be recovered through ruminants (Luke 2015).

Determining the value of feed protein in the feeding of ruminants is difficult due to the complexity of the digestive system. The most important thing to consider when evaluating protein is the feed protein that degrades in the rumen and the microbial protein that forms there. The protein absorbs in the small intestine expresses the amount of amino acid utilizes in the small intestine. Correspondingly, the balance of rumen is used to indicate the adequacy of nitrogen to the needs of rumen microbes (Luke 2015).

The interchangeability between silage and concentrate feed varies depending on the study, the digestibility of the silage and the ratio of the concentrated feed to the silage. The dry matter intake of silage decreased by 0,64 kg per kilo of concentrated feed (Ettala & Lampila 1978). According to a summary of 121 studies made at 1996, the average dry matter substitution ratio for silage was 0,39 with a standard deviation of 0,3 (Ryhänen, Huhtanen, Jaakkola & Ahvenjärvi 1996). Following to 16 comparison made by the Agrifood Research Finland, the ratio was 0,53 with a standard deviation of 0,12. In this case, the average daily amount of concentrated feed of cow was tree to 15 kg (Huhtanen 1998). As the amount of concentrate feed, the digestibility of the silage and the fermentation quality of the silage increases, the silage substitution rate increases. The amount of concentrated feed less than ten kilograms the substitution ratio was 0,51 and at higher than ten kilograms the ratio was 0,61 (Huhtanen 1998). According to this study, on average, one kilogram of concentrate feed in required to compensate for the 10 g/kg decrease in D-value of silage.

The feeding experiment conducted at the Agrifood Research Finland in 1998 compared the effect of primary and regrowth grass silage harvesting with protein supplementation on milk yield (Khalili, Sairanen, Nousiainen & Huhtanen 2005). The milk productivity was significantly more effective with primary grass than regrowth grass, although the utilization of metabolizable energy and absorbed amino acid of primary grass was lower than regrowth grass. The article explained the difference with higher nitrogen content of primary grass, which caused a high milk urea content. The nitrogen utilization decreased, and milk production increased when increasing the rapeseed cake concentration, for explanation provides very high nitrogen content of primary grass silage (Khalili et al. 2005). Another key point for considering the quality of grass silage can be found in the article written by Alstrup, Søegaard and Weisbjerg (2016), which explored the importance of harvesting season and maturity for the forage-to-concentrate ratio, as an effect to digestibility of the feeding. According to their feeding experiment conducted at Aarhus University in Denmark in 2009, the early maturity cuts increased the intake of dry matter and energy, the yield of energy-corrected milk and milk protein concentration and decreased fat concentration. For the conclusion, article emphasised digestibility rather than forage-to-concentrate ratio or maturity.

The milk yield effect of clover silage and ryegrass silage studied in 2015 at Aarhus University in Denmark. Johansen, Søegaard, Lund and Weisbjerg found that silage containing 50 % of either red or white clover increased the energy-corrected milk yield by 2,3 kg compared to the ryegrass silage. When comparing, white clover had the better milk yield than red clover (Johansen et el. 2017).

McNamara et al. (2003) investigated the effect of rumen-protected fat on cow fertility and milk yield. Experiments with two different preparations did not show a positive effect of the rumen-protected fat on the pregnancy of cows at the end of the feeding test. Both preparations increased milk yield during the first 12 weeks of lactation period, but decreased protein content during the first six weeks. One product had a significant positive effect on milk yield throughout the lactation period, but, the other did not. The contradictory result was explained by the end of feeding test before the end of lactation period. This Irish study used 201 Holstein Friesian spring-calving cows with the average parturition ratio of 2,9 and the forage-based feeding.

According to the silage feeding test conducted at the Agrifood Research Finland, Maaninka in 2008, a silage with D-value of 690 g/kgka produced an average 31 kg milk per day. Correspondingly, at a D-value of 650 g/kgka cows produced an average 28,8 kg milk per day. There were 42 Holstein cows in the experiment, of which six were calving first time. Three protein contents used in this study were 14, 18 and 21 % of concentrated feed. Decreased silage digestibility by 10 g/kg reduced daily protein yield by 18,2 g and fat yield by 24,3 g. Increasing the amount of concentrated feed per kilo of dry matter increased the protein yield by 30,3 g per day and the fat yield by 24,9 g per day (Juutinen 2011).

Vauhkonen (2011) studied the importance of silage harvesting time and times for the economy of dairy farm in her master’s thesis. The conclusion was that the economic differences between the silage harvesting strategies were rather small. Early harvested silage with an average D-value of 691 g/kgka gave the best result in optimal milk production. In the three-harvest strategy, the digestibility of silage was extremely high, with a high impact on milk yield and a low proportion of concentrated feed. On the other hand, the cost of harvesting rises with three rounds of harvesting, which would require high grain prices to be profitable.

3 Theory

Profit can be defined as the difference between returns and costs. If the return is greater than the cost of production, production is profitable. Mathematically, profit can be determined (Varian 1999, 327):

$$\pi = \sum_{i=1}^{n} p_i \, y_i - \sum_{i=1}^{m} w_i \, x_i \qquad (1)$$

In this equation, π represents profit, the first term describes the return on products and the next one the cost of production. Suppose a company, in this case a dairy farm, produces n units of products (y_1,…, y_n) and uses m units of inputs (x_1,…, x_m). Product prices are (p_1,…, p_n) and input prices are (w_1,…, w_m). According to Coelli, Rao, O´Donnell and Battese (2005), technical efficiency can be measured in either output or input orientation. This study focused on technical efficiency in output orientation.

3.1 Frontier production function

Production is efficient when operating at the upper limit of the set of production, i.e., the production curve (function):

$$y = f(x) \qquad (2)$$

The production function expresses the highest output at each level of input usage.

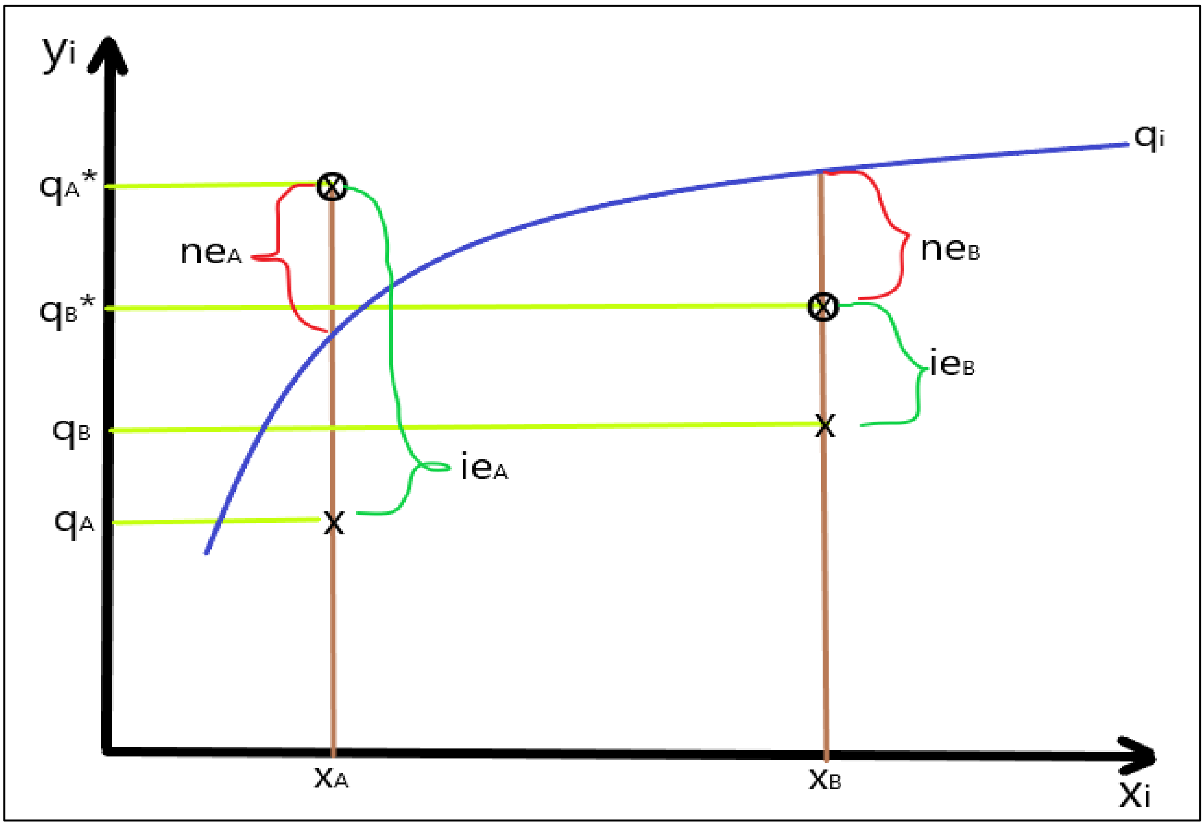


Figure 2. Stochastic frontier production function (Coelli et al. 2005, 244).

Figure 2 shows a stochastic production front function where the deterministic production function, e.g., $q_i = \exp(\beta_0 + \beta_i lnx_i)$ depicts an efficient production frontier with decreasing returns to scale. The inputs by two different companies (in this case, dairy farms) are plotted on the horizontal axis x_i and the output on the stakes on the vertical axis y_i. Dairy farm A uses inputs at level x_A to produce q_A and Dairy farm B uses inputs x_B to produce q_B. These observed values are indicated in Figure 2 by the exes. The output quantities produced without inefficiency would be at the following levels. (Coelli et al. 2005, 243):

$$q_A^* \equiv (\beta_0 + \beta_i lnx_A + v_A) \qquad u_A = 0 \qquad (3)$$

$$q_B^* \equiv (\beta_0 + \beta_i lnx_B + v_B) \qquad u_B = 0 \qquad (4)$$

In Figure 2, the values q_A^* and q_B^* according to the efficiency frontier are indicated by circled exes. Similarly, the inefficiency error u_A is denoted by a green parenthesis ie_A and u_B is denoted by a green parenthesis ie_B. The value q_A^* of the dairy farm A is above the frontier because the normal error term v_A is denoted by a red parenthesis ne_A is positive. The corresponding value q_B^* of the dairy farm

B is below the efficiency frontier because its normal error v_B, denoted by its red parenthesis ne_B is negative (Coelli et al. 2005, 243).

This model can, at least in principle, distinguish the errors caused by random variation and inefficiency (Kumbhakar & Lovell 2000, 72). The stochastic frontier production function, in which the error term is composed as described above, was introduced by Aigner, Lovell and Schmidt in 1977 and by Meeusen and van deer Broek in 1977.

$$lny = lnf(x) - u + v \qquad (u \geq 0) \qquad (5)$$

In Shephard's output orientation technical efficiency is defined as the ratio of observed output to the output defined by the efficiency frontier (Coelli et al. 2005, 244).

$$TE = \frac{y}{f(x)e^{v}} = \frac{f(x)e^{-u}e^{v}}{f(x)e^{v}} = e^{-u} \qquad (6)$$

The econometric estimation of the maximum likehood estimation of a stochastic production frontier function model requires the following requirements for the distribution of error terms: the statistical normal error term v follows the normal distribution and the mean of the terms is zero and the constant variance is σ_v^2. Similarly, the inefficiency term u gets the positive values of the normal distribution or the values of the normal distribution minus the negative values. The inefficiency term u is determined by the standard scale parameter σ_u^2 (variance). The terms are independently determined. (Greene 1980.)

$$v \sim N(0, \sigma_v^2) \qquad (7)$$
$$u \sim N^{+}(\mu, \sigma_u^2) \qquad (8)$$

When u obtains positive values of the normal distribution $\mu = 0$ (positive normal distribution) and $\mu \neq 0$ (truncated positive normal distribution). These assumptions result in a left-skewed distribution for the total error terms $\varepsilon = -u + v$. It is rare for statistics to have significantly large positive residual terms. The more common are the remarkably large negative residuals. (Henningsen 2019.)

When examining the inefficiency of the production function, the variance of the inefficiency term Var(u), denoted by the parameter $\gamma = {\sigma_u^2}/{\sigma^2}$, is used as an indicator. The parameter γ gets values between zero and one indicating the need for an inefficiency term: An inefficiency term is necessary if parameter γ is close to one, if $\gamma = 1$ then a normal error term is unnecessary and the standard deviation can be fully associated with inefficiency (Jondrow, Lovell, Materov & Schmid 1982). Similarly, if $\gamma = 0$, then the deviation can be explained by a normal error and the result of the stochastic frontier production function corresponds to the least squares function (OLS). The variance of the inefficiency term is calculated by the equation (Hennigsen 2019, 243.)

$$Var(u) = \sigma_u^2 \left[1 - \frac{\frac{\mu}{\sigma_u} \phi\left(\frac{\mu}{\sigma_u}\right)}{\Phi\left(\frac{\mu}{\sigma_u}\right)} - \left(\frac{\phi\left(\frac{\mu}{\sigma_u}\right)}{\Phi\left(\frac{\mu}{\sigma_u}\right)} \right)^2 \right] \qquad (9)$$

In the equation $\Phi(.)$ denotes the cumulative distribution function and $\phi(.)$ denotes the probability density function from the normal distribution. When u follows a positive normal distribution with $\mu = 0$, the equation can be converted to (Henningsen 2019, 243)

$$Var(u) = \sigma_u^2 \left[1 - \left(2\phi(0)\right)^2 \right] \qquad (10)$$

3.2 Marginal products

For calculating the marginal product on inputs, the function f(x) is assumed to be twice continuous differentiable. The marginal product MP_i describes the ratio of input i in the production y, where the highest possible output of y can be achieved by increasing input x_i while other inputs remain constant. Thus, it is a partial derivate of the function f(x) with respect to the input x_i. (Gravelle & Rees 1992, 181.)

$$MP_i = \frac{\partial f_i(x_1, x_2)}{\partial x_i} \qquad (11)$$

Considering the coefficient type definition for the stochastic frontier production model and assuming a random error v is zero, the marginal product is scaled down multiplied by the level of technical efficiency.

$$\frac{\partial y}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} e^{-u} = TE \frac{\partial f(x)}{\partial x_i} = TE\alpha_i \frac{f(x)}{x_i} \qquad (12)$$

However, technical efficiency does not affect output elasticities or scale returns (Henningsen 2019, 241)

$$\frac{\partial y}{\partial x_i} \frac{x_i}{y} = \frac{\partial f(x)}{\partial x_i} e^{-u} \frac{x_i}{f(x)e^{-u}} = \frac{\partial f(x)}{\partial x_i} \frac{x_i}{f(x)} = \frac{\partial lnf(x)}{\partial lnx_i} = \alpha_i \qquad (13)$$

Derivate translog production function

$$lny = \alpha_0 + \sum_i \alpha_i ln\, x_i + \frac{1}{2} \sum_i \sum_j \alpha_{ij}\, lnx_i lnx_j \qquad (14)$$

regarding the quantities of inputs:

$$MP_i = \frac{y}{x_i} \left(\alpha_i + \sum_j \alpha_{ij} lnx_j) \right) \qquad (15)$$

The marginal production of all inputs need not be continuously positive, for example, the continuous increase in crude protein per cow will, after a certain limit, reduce milk yield. This is due to acidification of the rumen caused by excessive concentration of the protein of diet. Positive marginal production on other inputs, in this case, energy and fit value, make the result positive. The assumption is that there is always a positive marginal production on one input. The requirement for a input set X(y^0) at the output level y^0 is that the output of the input set combination is at least y^0. (Gravelle & Rees 1992, 181.)

$$X(y^0) = \{x | f(x) \geq y^0\} \tag{16}$$

3.3 Technological indicators

In addition to input variables, a production function based on input and output observations may have other variables affecting output rates. The inclusion of these factors in the applied production function is necessary when they influence the production process (Huang & Liu 1994; Lothgren 2000; Wang 2002). This data contains categorical variables such as feeding type or barn type, which may affect milk output. These environmental factors (z) can be added to the production function after first being converted to dummies.

$$y = f(x, z) \tag{17}$$

This production function can be used to examine the relationship between the environmental factor (z) and the output. As a translog production function is used the following model (Huang & Liu 1994)

$$lny = \alpha_0 + \sum_i \alpha_i ln\, x_i + \frac{1}{2}\sum_i \sum_j \alpha_{ij}\, lnx_i lnx_j + \alpha_z z \tag{18}$$

The ineffectiveness term of the translog production function is calculated as follows (Huang & Liu 1994)

$$lny - \ln(\eta) = \sum_i \alpha_i Z_i + \sum_i \sum_j \alpha_{ij} Z_i \ln(X_j) + W \qquad (19)$$

The output of SFA is described by the term η, y is the output of the dairy farm, the environmental factors are described by Z_i, the inputs are described by X_j and W is the random error. However, one should be cautious about the value of the inefficiency term and interpret it as a rough estimate of effect of a factor expressed by a dummy variable rather than an exact value (Henningsen 2019, 265). For example, if the loose-housing factor is positive, it means that it has a positive effect on output. On the other hand, the log likehood value can be used as a test when approximating the influence of dummy variable.

3.4 Production technology

Production technology sets the boundaries for a company by defining the possible relationships between inputs and outputs, that is, the relationship between the composition of inputs and output (Ryhänen & Sipiläinen 2018). Production know-how is also part of production technology, which can be defined as a combination of physical, biological and technical factors (Ryhänen & Sipiläinen 2018). Production technology is examined in this paper at three levels: the nutrient level of cows, the feed component level and the total number of cows and amount of forages. The main nutrients in feeding cattle are crude protein and energy. They do not substitute each other, so the output is determined by the minimum of available inputs. The second level consist of feed components, mainly silage and concentrated feed, which are largely composed of the same basic nutrients. Therefore, they can substitute each other. The third level consist of cows and total amount of forages.

In Leontief technology, i.e. fixed technology, the output is determined by the lower input (Varian 1999, 316)

$$y = f(x1, x2) = \min\{x1, x2\} \qquad (20)$$

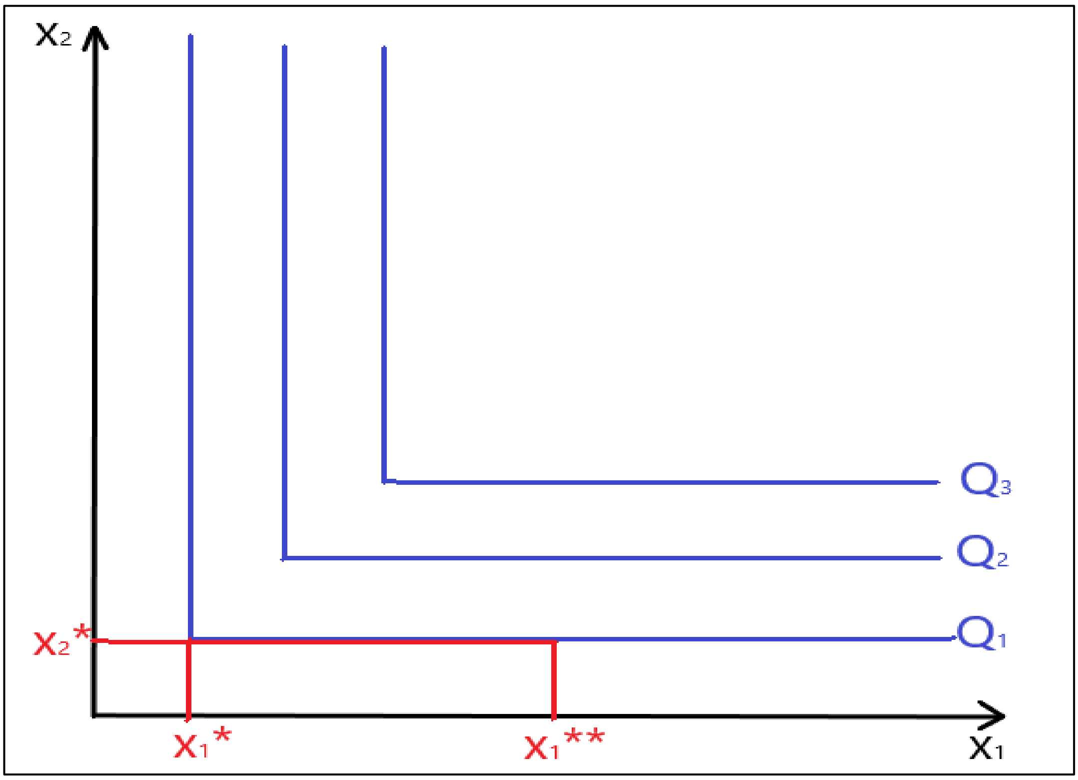


Figure 3. Leontief production technology.

In this first level production function, X_1 is the amount of crude protein and X_2 is the amount of energy. In Figure 3, the x-axis has the crude protein content (X_1) and the y-axis the total energy (X_2). Q_1, Q_2 and Q_3 are isoquants, one isoquant gives all input bundles that produce the same amount of output. Figure 3 illustrates that only increasing the amount of crude protein from level X_1* to level X_1** does not increase output when the amount of energy is at level X_2*. The amount of energy acts as the minimum input, which determines the output.

Cobb-Douglas (CD) production technology describes situations in which inputs are interchangeable (Varian 1999, 317)

$$y = f(x_1, x_2) = Ax_1^{\alpha}x_2^{\beta} \tag{21}$$

The CD production function has a standard factor A that can be used to scale the production. The variables x_1 and x_2 represent the inputs, while the parameters α and β indicate how much the output changes when the input is changed. The sum of the parameters α and β indicates the magnitude of the scale output. In CD production function, there is a constant elasticity of substitution 1. More flexible forms of function can be used in the analysis of second level production technology, which compares the effect of feed components used as inputs on output. The main components of feed are silage and concentrate feed, which can substitute each other (Huhtanen, Rinne & Nousiainen 2008). Both contain varying amounts of energy and crude protein. It may be also possible that protein and energy can marginally substitute each other. The excessive protein input may be converted to energy and the excessive quantity of energy may be improve the efficiency of protein utilization. On the other hand, cows can produce milk 20 litres per day with supplies of tissues (LUKE 2015, 54).

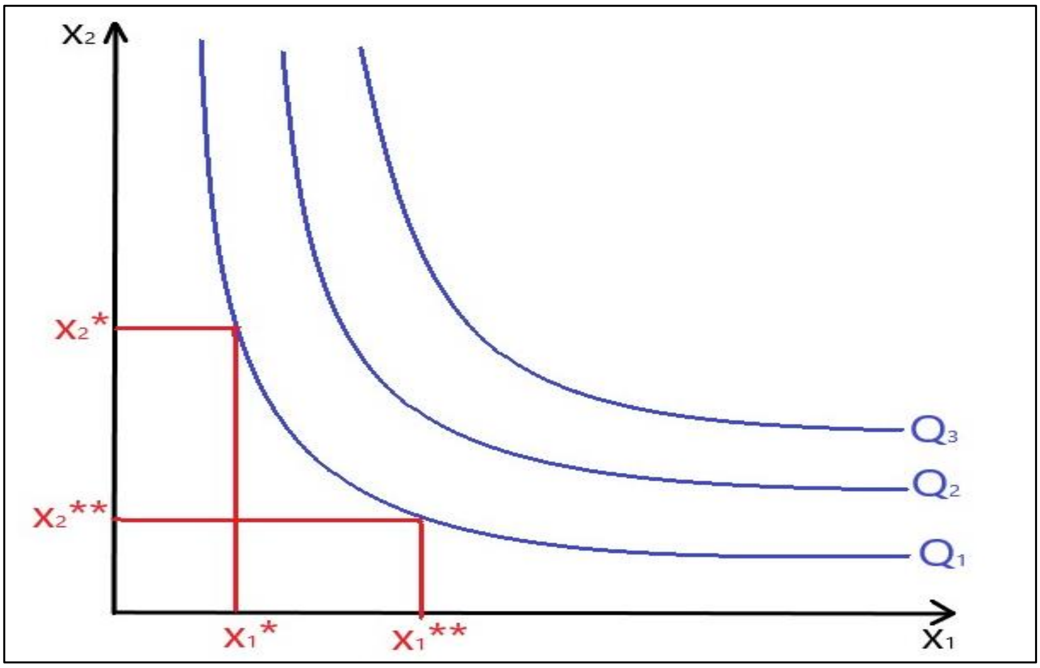


Figure 4. Cobb-Douglas technology.

Two alternative inputs of the second level production function are shown in Figure 4. In the first alternative, input amounts X_1* and X_2* are used, in the second alternative X_1** and X_2** respectively. Both technology options produce the same output on isoquant Q_1. Thus, it can be concluded that the inputs X_1 and X_2 are at least partially substitutable with each other.

The third level looks at total forage consumption and cow numbers. At this point, we return to Leontief production technology when it comes to the cow's maximum ability to eat. Forage use is limited by the cows' limited dry matter intake, which is three to four per cent of the body weight of a regular cow (Aspila et al. 1994, 49). This means a maximum of 24 kg of dry matter per day for a 600 kg weighted cow. In order to achieve technical efficiency, the cow's eating capacity must be fully utilized during the lactation period, while it is not worth feeding over the cow's current need during the dry period. Cobb-Douglas technology can be applied below the cow's eating capacity. The issue of total nutrient and cow balance is discussed in Section 2.3.

Summarizing these levels, the only way to effectively influence the economic result of milk production through production technology is the optimum supplementation of silage with concentrated feed. Nutrients cannot be freely combined because they are poor sobstitutes as substitutes for one another. Similarly, cows are limited in their ability to eat and nutrients can be limitedly enriched in their forage, because their digestive system requires enough fiber to function (Aspila et al. 1994, 51). However, the price of the nutrients in the diet depends on the component they are included in the diet. At a general level, it is conceivable that in concentrated feed (cereals and rapeseed) the cost of producing nutrients is in principle higher than in silage (ProAgria 2018).

3.5 Production risks in milk production

Milk production differs from regular industrial production processes due to varying production conditions. Keeping ambient conditions optimal for production is challenging and not all risk factors can be excluded. This also causes problems in the practical application of research results, since research conditions may not correspond to actual production conditions (Hardaker, Anderson & Huirne 1997, 11). Only the risk factors that appear in the milk production chain from field to milk tank of farm are discussed here. The thesis focuses only on production risks (Hardaker et al. 1997, 6; Olson 2004, 294).

The first risk is the weather risk in forage production when looking at the milk production chain. Especially the study year 2018 was exceptionally dry and hot (FMI 2018). This caused problems in silage production, which requires enough rainfall throughout the growing season. As regards clover grasses, the statistics show that the problem of 2018 was exacerbated by the wetness of the previous year (FMI 2017).

Losses in storage are the next potential risk factor in the production chain. No statistical losses can be deduced from storage, as the quantities of forage produced and used are not recorded. In the case of silage analyses, the statistical material does not indicate preserved quality either.

The realized risks of feeding are best visible in the statistical data. The same cow-based feeding results in different production results, leading to different revenues and milk yields. This thesis aims to explore the factors influencing milk yield.

Some of the risks realized in the barn may be due to the risk of disease instead of feeding. Either as a result of feeding from contaminated forage (salmonella) or from people visiting the farm or from animals (salmonella and viral diarrhoea). Realization of the disease risk often results in quite a significant drop in milk production.

4 Research material and methods

4.1 Research material

The material used in the thesis in a sample of the MoC of milk production for 2018 of ProAgria, compiled from the dairy farms of Valio. The original statistics consist of 3 522 MoC:s, of which 1948 relate to 2 829 forage analysis (Artturi-analysis). The study uses these 1 948 MoC:s, because forage analysis is required to answer the research questions. MoC:s includes 91 984 individual observations of cows. In farms with more than one MoC, some of the cow observations are directed to the same cows. In these cases, the interval between observations is usually about six months, with about 20 % of the cows disappearing because of a median calving of 2,5 per cow. Correspondingly, due to the lactation period of cows of about ten months and the dry period of two months, approximately 30 % of cows in production are replaced annually. The remaining 50 % of cows are at different phase of lactation period and it was therefore not considered appropriate to delete the MoC:s. These with more than one forage analysis have used average forage analysis. In most cases, the use of bale silage forage is the explanatory factor for several forage analyses for the same calculation.

The farms are numbered sequentially from number one to match MoC:s for the same farm. They are numbered accordingly for inclusion in forage analysis. The calculations are missing farm identifiers and cannot be associated with the original farms. The original MoC includes milk amount, fat percentage, protein percentage, cows, average total dry matter intake, percentage of concentrated feed, total protein in the diet, total diet OIV[14], total diet ME corrected. New variables have been calculated based on the data in the calculation average milk yield per cow, total livestock dry matter intake, average silage and concentrated feed intake, approximation of concentrated feed cost and milk yield per cow. The variables that classify the original file are coded according to tables one and two for the MoC.

[14] OIV: The protein absorbs in the small intestine expresses the amount of amino acid utilizes in the small intestine.

Table 1. Categorical variables of livestock and feeding.

Code	Type of livestock	% and (number)	Code	Type of feeding	% and (number)
0	Not known	**1,8 (35)**	**0**	Not known	**2,2 (42)**
1	Tied-housing system	**43,4 (843)**	**1**	Separate feeding	**68,6 (1333)**
2	Loose-housing system	**54,1 (1052)**	**2**	Mixed feeding	**12,6 (244)**
3	Other	**0,7 (13)**	**3**	Mixed feeding and supplement	**16,7 (324)**

Table 2. Categorical variables of sharing mode of concentrated feed and milker

Code	Sharing mode of concentrated feed	% and (number)	Code	Type of milker	% and (number)
0	Not known	**5,6 (109)**	**0**	Not known	**0,8 (16)**
1	Other	**0,6 (11)**	**1**	AMS	**28,2 (547)**
2	Mixing wagon	**7,9 (153)**	**2**	Milking parlor	**26,4 (510)**
3	Milking parlor and kiosk/AMS and kiosk	**4,7 (92)**	**3**	AMS and milking parlor	**0,5 (9)**
4	Milking parlor/AMS	**13,7 (266)**	**8**	Pipeline milking machine	**44,3 (861)**
5	Kiosk	**22,8 (443)**	**9**	Bucket milking machine	**0 (0)**
6	Rail carriage	**28,5 (554)**			
7	Hand share	**16,2 (315)**			

Silage forage analysis includes livestock and the number of MoC, forage code, forage name, which can also be used to infer silage harvest round, crude protein of silage, legume content, D-value, corrected ME, OIV, and calcium and phosphorus levels if analysed.

A variable has also been added to the data base to illustrate the impact of the protein content of silage on the cost of concentrated feed. The data on the quantity of concentrated feed and the crude protein content of silage are used to calculate the cost of concentrated feed (Appendix 1).

$$w_v = \frac{((V_d - V_s * (1 - C_v)/C_v)) - V_v) * (P_r - P_v)}{(V_r - V_v) * D_v} + P_v \qquad (22)$$

In the equation w_v is the price of concentrated feed, which is determined by the protein content V_d of the whole diet, the crude protein content V_s of the silage and the protein content of the concentrated feed C_v as described above. The price of concentrated feed is proportional to that of wheat and rapeseed, estimated at EUR 160/ tonne for wheat and EUR 300/tonne for rapeseed. Correspondingly, the crude protein content V_v of wheat is estimated at 150 and that of rapeseed V_r 340. The dry matter content D_v of the concentrated feed is estimated at 0,86. The price of the concentrated feed is calculated by comparing the protein content of concentrated feed required and prices announced before. The cost of concentrated feed calculated by multiplying the price with the amount of it which is calculated from the data.

$$W_v = \frac{A_v * w_v}{1000} \qquad (23)$$

The total cost of concentrated feed W_v is determined by the amount of concentrated feed A_v and the price w_v as described above.

4.2 Determining productivity through quality-adjusted output

The producer price of milk is determined by the hygienic quality of the milk and the fat and protein content. Based on hygienic quality, milk is divided into three classes E, 1 and 2. E-class milk is paid a liability premium to contracting dairy farms. More than 97 % of the milk received by Valio is included in this category. There is no additional price for milk category 1 but there is a quality reduction for category 2. However, the protein and fat content of the milk has a more significant effect on the milk price paid to producer, with a difference of 39,2 % between the lowest and the highest producer price.

In this paper, the technical efficiency is examined based on feed inputs and the return of milk obtained from them. The data provide accurate information for determining the return. Milk return can be derived from the amount of milk and the protein and fat content that influence the price. The hygienic quality is not stated in the statistic and it is not considered in the pricing. The milk return is determined by the base price of the Cooperative Tuottajain Maito and concentration adjustments for 2018. There is no monitoring-based price for inputs.

4.3 Quantitative research

Appendix 2 groups the material in this thesis according to a commonly used model. Farms were numbered consecutively from one onwards, so the farms can be distinguished from each other. But based on the numbering, the data could not be attached to the original farms. Similarly, silage analyses and MoC:s have a numerical identifier that allows them to be combined. These tags allowed the aggregation of descriptive statistics from a single MoC.

The endogenous variable in the function was milk return. The average amount of milk per MoC and per cow were determined based on the number of cows and milk amount in the original data base. The average milk return per cow per calculation was formed by multiplying the amount of milk by the price of milk formed based on protein and fat percentages. An efficiency analysis based on milk volume alone was not considered enough, as the price of milk calculated on the data base varied by 39,2 %. The price was calculated according to Cooperative Tuottajain Maito pricing criteria, and the price does not include possible subsidies.

The energy and protein amounts of the feed and the adapter value, which was defined in detail in appendix 1, have been used as exogenous variables. The amount of energy was calculated based on the total amount of feed and the energy content. Correspondingly, the amount of protein was calculated based on the total amount of feed and the crude protein content. Energy and protein were the main components in the feed and can be reliably determined from the data base used. The purpose of the fitting variable has been to bring additional information to the production function by determining the price of concentrated feed based on the required protein content and multiplying it by the amount of concentrated feed.

The clover content of silage, the type of barn, milking machine and feeding, and the method of share of concentrated feed have been used as explanatory variables. The bound for the clover content of silage has been set at 20 %, as this would make the differences between it and grass silage more apparent. The types of barns in the study were tied-housing and loose-housing. Among the milking system types in the comparison were automatic milking system (AMS), milking parlor, AMS and milking parlor, pipeline milking and bucket milking. Of the feed share methods in the comparison were mixing wagon, share by milking parlor or robot supplemented with kiosks, share by milking parlor or robot alone, kiosks, rail carriage and hand share.

4.4 Methodology

The R-Studio is used as an analysis software because of it is freely available, and it is easy to modify as required by the analyses. R software like “frontier”, “micEcon”, “lmtest”, “plm”, and “rgl” have also been used in the analyses. Figure 11 has made by SPSS. The statistical data has been fed through Excel spreadsheet software, where the data have been also partially edited. For example, MoC:s to which forage analyses could not be linked have been removed from the data. At the same time, new categorization and dummy variables have been added to the data as well as new variables transformed from the original variables.

The quantitative statistical analysis examined the technical efficiency of milk production. Statistical testing led to a flexible logarithmic stochastic frontier production function. The translog function is based on the interchangeability of the inputs, but in this case the total amount of protein of daily diet and the amount of energy are basically not interchangeable. During the estimation, the error term is composed into noise term and inefficiency:

$$lny = \alpha_0 + \sum_i \alpha_i \, lnx_i + \frac{1}{2}\sum_i \sum_j \alpha_{ij} \, lnx_i lnx_j - u + v; \; \alpha_{ij} = \alpha_{ji} \qquad (24)$$

5 Results

This chapter summarizes the main results. The more detailed results are presented in Appendix 3. The analytical models were compared to each other using the likehood ratio test. The production function used in the thesis, the stochastic frontier production function using the logarithmic interaction model, is as follows:

$$lny = \alpha_0 + \alpha_1 lnx_1 + \alpha_2 lnx_2 + \alpha_3 lnx_3 + \alpha_4 (lnx_1)^2 + \alpha_5 (lnx_2)^2 + \alpha_6 (lnx_3)^2 + \alpha_7 (lnx_1 * lnx_2) + \alpha_8 (lnx_3 * lnx_2) + \alpha_9 (lnx_1 * lnx_3) - u + v \qquad (25)$$

y is the average milk yield per cow, formed by multiplying the price per litre corrected for fat and protein content of milk by the average milk amount. The total protein amount of the daily diet is represented by the variable x_1 and the total energy x_2. The fitting value x_3 illustrates the cost variable derived from the protein content of concentrated feed calculated from relation between silage and diet protein content and multiplied by amount of concentrated feed (Appendix 1).

Table 3. The table shows the coefficients of the above equation with the accompanying data.

	Estimate Std.	Error	z value	Pr(>|z|)
α_0	-5,2064e+01	-21,4930e+01	-21,4930	<2,2e-16 ***
α_1	-2,1718e+00	1,7970e+00	-0,0012	0,9990357
α_2	1,9210e+01	1,1090e+00	17,3217	<2,2e-16 ***
α_3	-2,0616e+00	8,7969e-01	-2,3435	0,0191015 *
α_4	-1,2527e+00	4,6416e-01	-2,6988	0,0069585**
α_5	-3,4469e+00	2,6977e-01	-12,7775	<2,2e-16 **
α_6	3,9720e-02	1,7691e-02	2,2452	0,0247572 .
α_7	2,9489e-01	4,1512e-01	0,7104	0,4774748
α_8	2,9676e-01	1,8145e-01	1,6355	0,1019380
α_9	3,3440e-01	1,0019e-01	3,3376	0,0008451 ***
sigmaSq	2,6913e-02	3,7114e-03	7,2513	< 2,2e-16 ***
gamma	8,3938e-01	2,8199e-02	29,7657	< 2,2e-16 ***

Model 1: OLS (no inefficiency)
Model 2: Error Components Frontier (ECF)

```
  #Df LogLik Df  Chisq Pr(>Chisq)
1  11 2063.7
2  13 2084.7  2 42.148  2.183e-10 ***
```

Both the likehood ratio test and the gamma (γ) value of 0,84 indicate that the inefficiency term is significant. The coefficients α_2, α_3, α_4, α_5, α_6 ja α_9 are significant at the level of at least 0,1 (p-values marked in green). The coefficients α_1, α_7 ja α_8 are not significant at the level 0,1 (p-values marked in red).

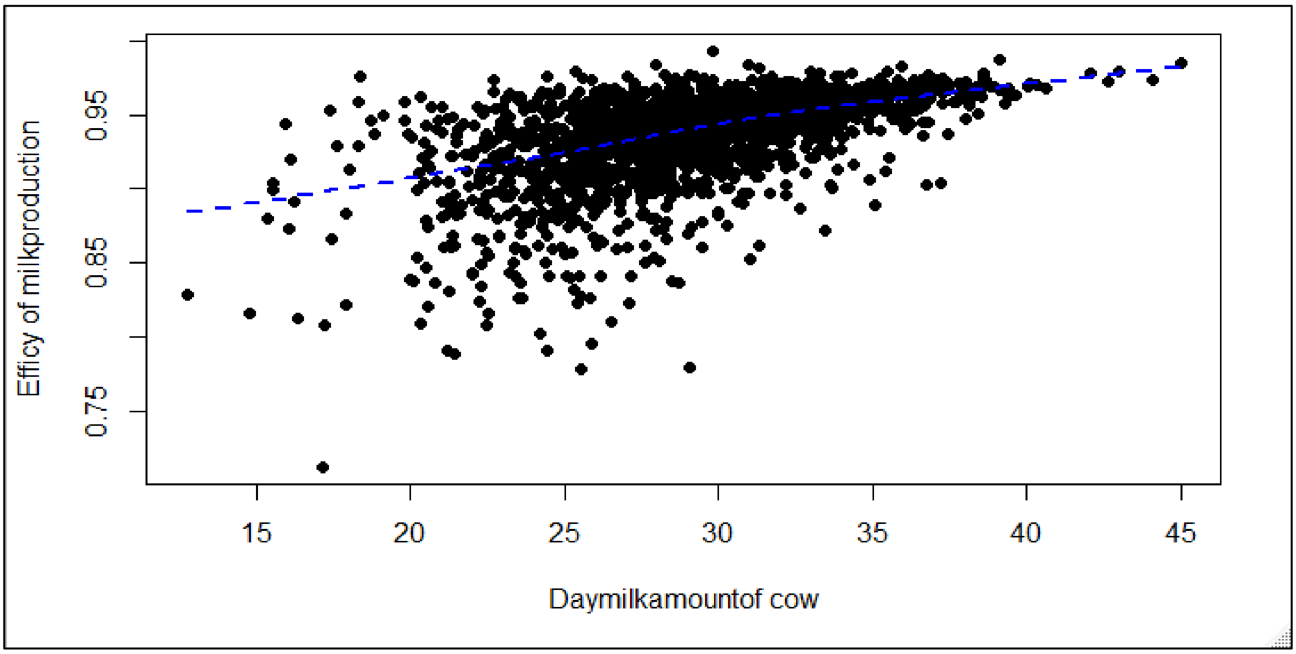


Figure 5. Comparison of MoC:s between milk production efficiency and daily milk amount.

In Figure 5, the y-axis has the milk production efficiency calculated by translog function. The x-axis shows the average daily milk amount in litres per cow. The mean of efficiency increased when the average amount of milk production of cattle increased. Nevertheless, the efficiency of 0,96 was reached at the level of 18 litres per cow in average. The variation of efficiency was decreased when average amount of milk increased.

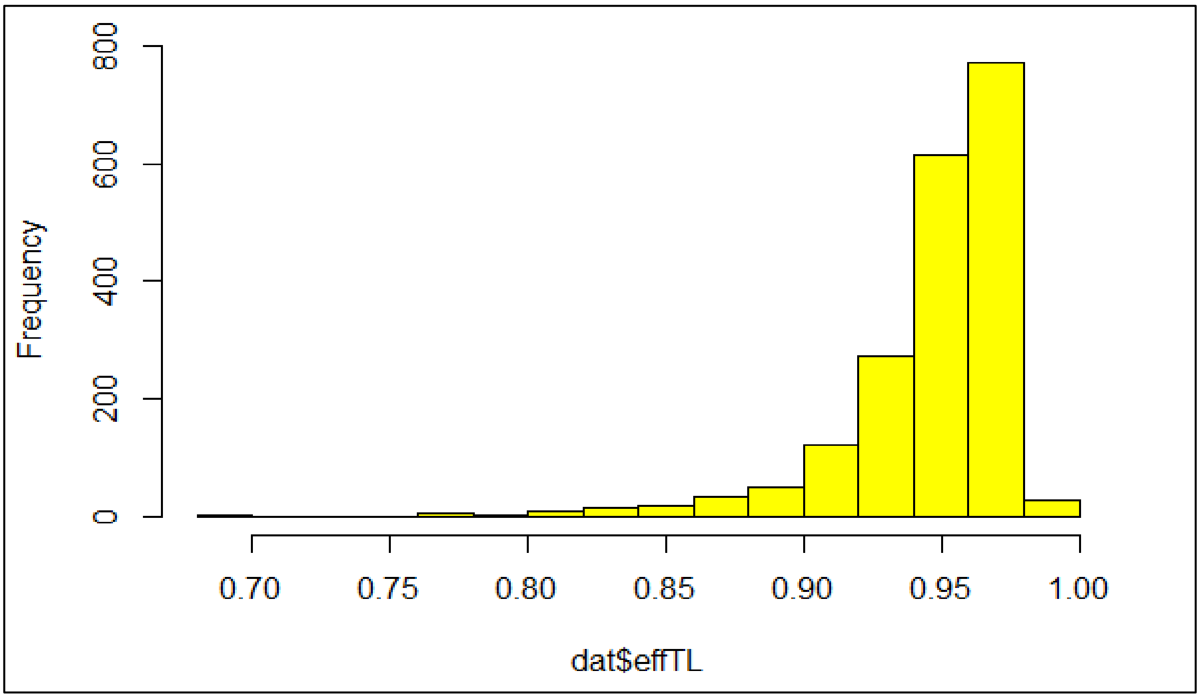


Figure 6. Distribution of technical efficiency.

Figure 6 shows the distribution of technical efficiency of MoC:s, calculated by translog function. In y-axis is the frequency of MoC:s and in y-axis is the technical efficiency categorised at five percent steps from 0,65 to 1,00. The majority of MoC:s achieved at least the efficiency of 0,95.

5.1 The responses of energy and protein

To determine the responses of energy and protein, a new translog function was formed by removing the fitting value. Other inputs are illustrated as before, total amount of protein of daily diet x_1 and total energy of daily diet x_2. The more detailed results are presented in Appendix 4.

$$lny = \alpha_0 + \alpha_1 lnx_1 + \alpha_2 lnx_2 + \alpha_3 (lnx_1)^2 + \alpha_4 (lnx_2)^2 + \alpha_5 (lnx_1 * lnx_2) - u + v \qquad (26)$$

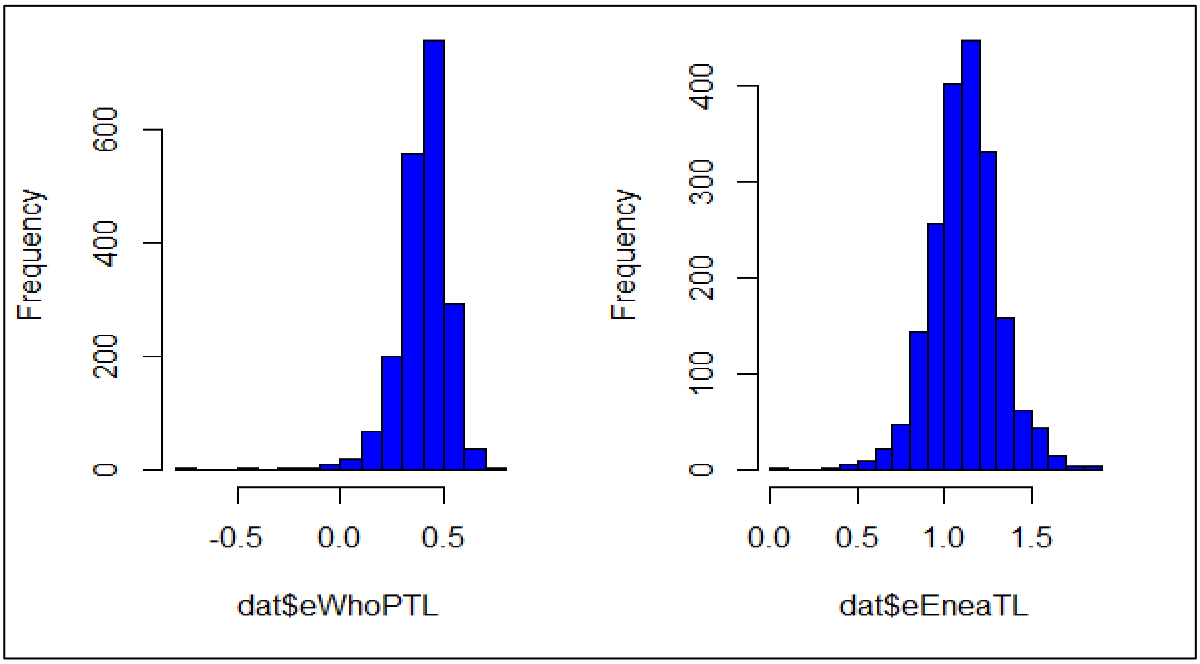


Figure 7. Elasticity of output by inputs.

Figure 7 illustrates the elasticities of output. On the right is the elasticity of output by protein and on the left is the elasticities of output by energy. In y-axis is the frequency of MoC:s and in y-axis is the elasticity of output by inputs categorised by 0,5 steps. Almost all elasticities of output of protein were positive and all ones of energy. This means, that it is profitable to increase amounts of protein and energy. As the marginal products, the output elasticities of inputs measure the marginal productivities. The changes of the input and output quantities are measured in relative terms; hence the units of measurement have no effect to output elasticities.

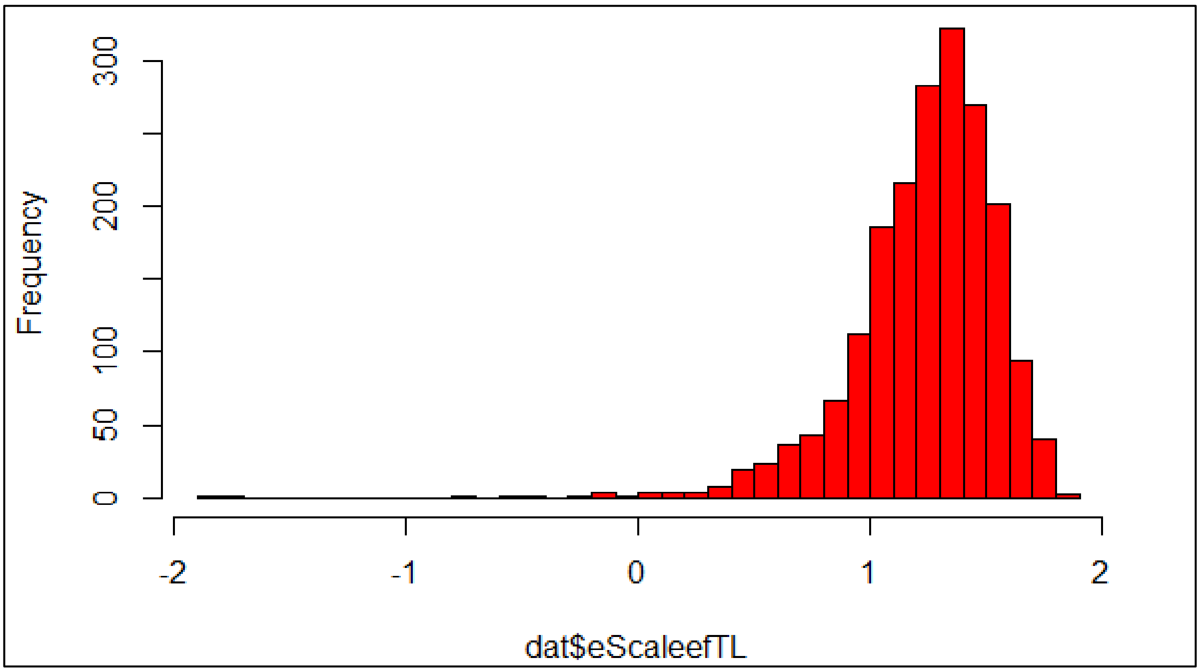


Figure 8. Elasticity of scale.

Figure 8 illustrates the returns of scale of the technology. In y-axis is the frequency of MoC:s and in x-axis is the elasticity of scale. The majority of MoC:s had increasing returns to scale more than one ($\varepsilon > 1$), which means that total factory productivity increases when increase the amounts of protein and energy.

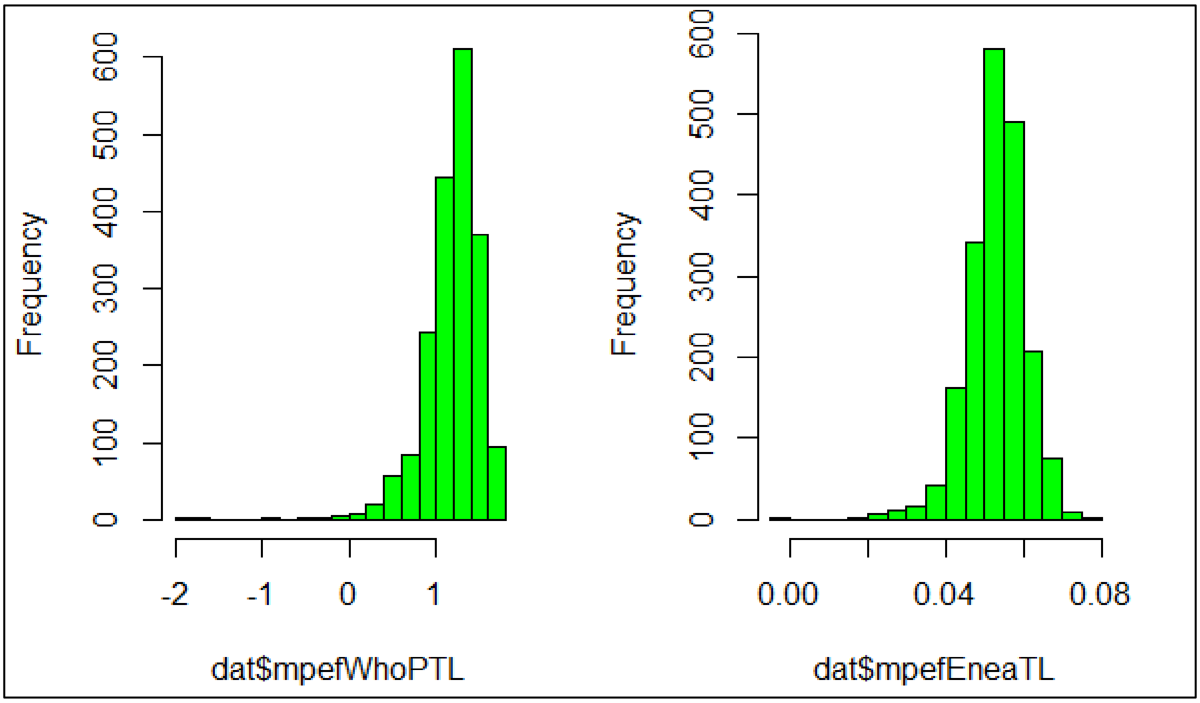


Figure 9. Marginal products of inputs.

Figure 9 illustrates the marginal products of inputs. On the right are the marginal products of protein and on the left are marginal products of energy. In y-axis is the frequency of MoC:s and in x-axis is the marginal products of inputs. In calculation has been used farm specific efficiency scores. Almost all marginal products of protein and all ones of energy were positive, which means that increase of these inputs were profitable.

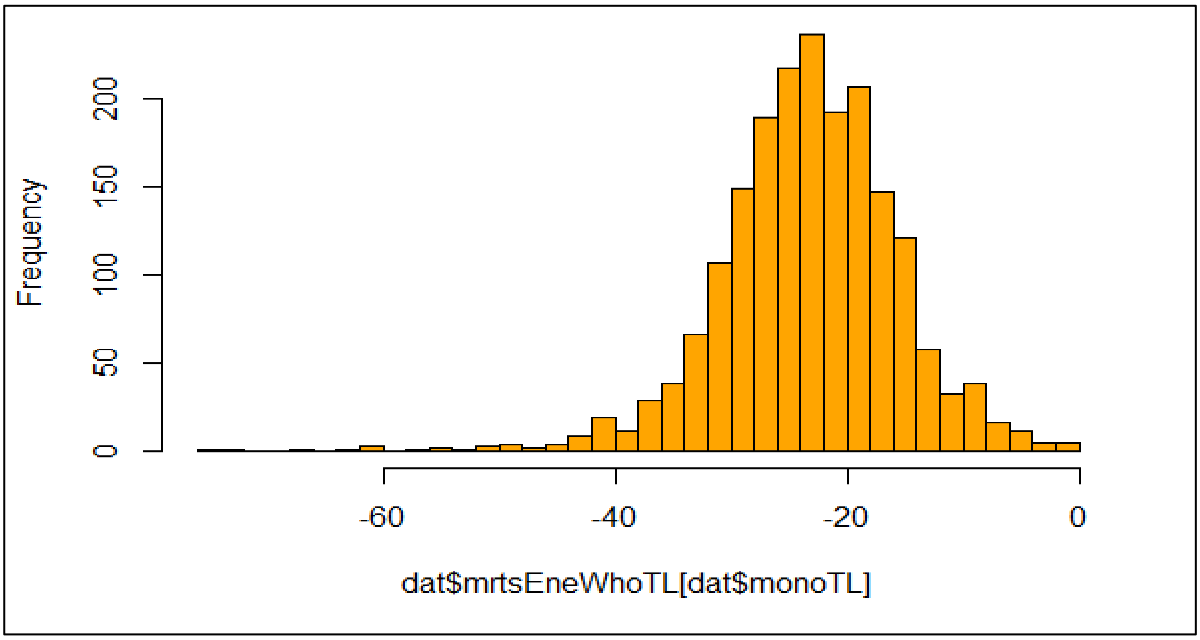


Figure 10. Marginal rate of technical substitution.

Figure 10 illustrates the marginal rate of technical substitution (MRTS) when the marginal product of protein was divided by marginal product of energy ($MRTS_{e,p}$). In y-axis is the frequency of MoC:s and in x-axis is the marginal rate of technical substitution. The mean of $MRTS_{e,p}$ was -22,976. This means that one unit of protein would substitute almost 23 units of energy.

The approximation of rate of prices for energy and protein was made as a linear regression analysis of prices and contents of components of concentrated feed. The cost minimizing price ratio was around 115. Cost-effectiveness would require that the absolute value of MRTS and the price ratio between inputs are equal. The more detailed results are presented in Appendix 5.

5.2 The effects of technological indicators

The initial stochastic frontier production function was used when considered the effects of technical indicators. Thus, the fitting value was added to function. Variables in the SFA logarithmic interaction model scaled to their mean values give a scaling yield of 0,19+0,94+0,02=1,15, by mean scaled translog function with confeedwholecost. The more detailed results are presented in Appendix 3. This indicates a positive scale advantage. The variance of u of 0,511 calculated from the Cobb-Douglas function shows that 51,1 % of the total error is due to inefficiency. 37,4 % of the total error in translog function is due to inefficiency. The effect of environmental factors on revenue was studied by including them in the translog function as a z variable. Other variables that were not replaced between analyses were the total amount of protein and the total energy for daily diet and fitting variable.

There were three feeding types in the comparison: separate, mixed and mixed supplemented with concentrated feed. In separate feeding, concentrate feed and silage are separated, in mixed feeding they are compounded and in a third alternative, a standardized mixed feed supplemented by the concentrated feed based on individual needs of the cows. The reference point was separate feeding. Of these, mixed feeding with supplement with concentrated feed proved to be best. However, it had a relatively low coefficient of 0,0082 with a p-value of 0,11. The mixed feeding coefficient was -0,00012 with p-value of 0,98. This means there was no statistically significant difference, if the reference level was 10 % risk.

The tied-housing was considered as a reference in the barn type comparison, then, loose-housing had a coefficient of 0,0055 with p-value of 0,15, which means that there is no significant difference. In the milker type comparison, the automatic milking system was the reference. The milking parlor had a coefficient of -0,0133 with a p-value of 0,011, a pipeline milking coefficient was -0,0091 with a p-value of 0,050, and the parlor and AMS coefficient were -0,053 with a p-value of 0,060, these results were statistically significant.

The reference level for concentrated feed sharing was the mixed feed option. Milking parlor or AMS supplemented with kiosk received a coefficient of 0,0093 with p-value 0,35, milking parlor or AMS without kiosk with 0,0130 and p-value 0,07, kiosk 0,0077 with p-value 0,23, rail carriage 0,0042 with p-value 0,50 and a hand sharing coefficient of 0,0092 with a p-value of 0,19. All these results were statistically insignificant.

In silage comparison, silage with a clover content of a least 20 % received coefficient of -0,018 with a p-value of 0,000, when the reference point was grass silage. When comparing silage harvest rounds, the summer harvest was best with a coefficient of 0,013 with a p-value of 0,002 and the autumn harvest had a coefficient of 0,011 with a p-value of 0,067, when the spring harvest was taken as the reference.

5.3 The approximations of the costs of feed and the rate of protein and energy

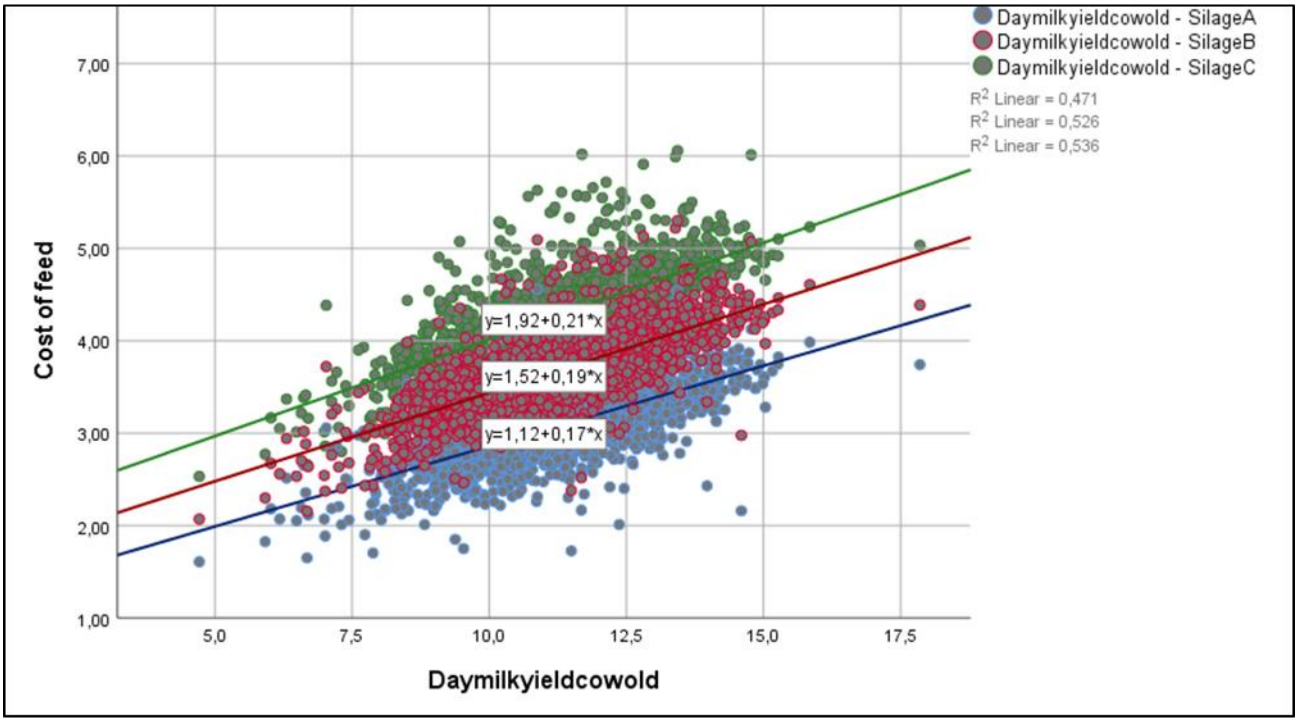


Figure 11. Costs of feed depending of the price of silage.

Figure 11 illustrates through an example calculation the effect of silage production cost on feeding costs. Fitted value has been used in the calculation, to which the production costs of silage have been added at three cost levels, 10, 15 and 20 cents per kilogram of dry matter. The lowest line (blue) describes the feed cost when the silage production cost was 10 cents, the middle line (red) when the silage production cost was 15 cents and the highest line (green) when the silage production cost was 20 cents per kilogram dry matter.

The approximation of rate between energy and protein gave the result of 64,5, when considered the subgroup made of MoC:s, which were at the level of 0,97 or more at efficiency. This is in the line with the rate of 65,3 by feeding plan of ProAgria. The plan based on LUKE's calculations of cow's energy and protein requirements. The more detailed results are presented in Appendix 5.

6 Discussion

Previous studies discussed in this thesis have mainly been based on limited numbers of cows and performed under controlled conditions. This has made it possible to eliminate interfering factors. Thus, the effect of a single variable has been easier to observe. On the other hand, relatively small samples easily cause distortions since one single observation may be significant for the result. This paper is based on data collected from functioning dairy farms. For this reason, it is problematic to determine reliably the variability of an individual variable. The effect of other variables cannot be excluded. Alternatively, the large data helps to prevent the error caused by individual observations. However, it should be noted that dairy farms must apply the research results to their practical activities, in which case it is appropriate to compare the effect of controlled and farm-level conditions with the analytical results.

Testing of the stochastic frontier production function began from Cobb-Douglas function, ending in a translog model based on the likehood ratio test. The model obtained was best suited to describe the statistics to be analysed. Of the individual variables α_1, α_7 and α_8 were not significant. The overall explanatory power of this model was the best of the alternatives studied. Returns of scale (1,15) indicated better productivity when increasing proportionally additional inputs. A one per cent increase in inputs given 1,15 % increase in output. The inefficiency term for Cobb-Douglas function explained 51,1 % of the total error and the inefficiency term for translog function 37,4 %. This was due to the better fit of the translog model with statistical material.

The model used in this thesis made it relatively easy to set up the MoC:s, or alternatively, the dairy farms in order of efficiency. Most of the input data were available from the MoC and supplementary data from the regional cooperative dairies. The fitting variable used in the calculations was not enough to describe the variation in input prices in the analysis, but it provided additional information on the effect of protein content of concentrated feed on feeding costs, and so, it improved the model fit. However, it could not describe the variation in input prices at the farm level, as the prices depended on differences in production costs, lot size, freight costs, etc.

Previous dissertations presented in this thesis focused on the elucidation of silage and concentrated feed as components in ruminant feeding. This dissertation was focused on elucidating protein and energy responses as well as the effects of environmental factories. According to the results of previous studies presented in this dissertation, a kilo of silage dry matter replaced 0.39 to 0.63 kilos of concentrate feed, depending on the quality of silage and the amount of concentrate (Ettala & Lampila 1978; Ryhänen et al. 1996; Huhtanen 1998). Alstrup et al. (2016) emphasized the effect of harvesting season and digestibility on productivity rather than the forage-to-concentrate ratio. When elucidating the responses of protein and energy, a new equation was formed in which the fitting value was removed from the inputs. The response of energy was 0,91 and the response of protein was 0,21. In this case, the scale efficiency became 0,21+0,91=1,12, which indicates a positive scale advantage. The mean of $MRTS_{e,p}$ was -22,976, based on that, one unit of protein would substitute almost 23 units of energy. The cost minimizing ratio is around 115. Therefore, these two results pointed that it was profitable to all farmers to increase the use of protein. The mean rate of energy and protein in the feed was 64,5, when considered MoC:s with efficiency of 0,97 or more. This was in the line with the requirements of cows calculated by LUKE (2015). The analysis of this thesis indicated that protein and energy would substitute each other in the diet, but the analysis did not take a count cows potential to use the supplies of tissues for production of milk (LUKE 2015, 54). According LUKE (2015), cows can produce milk 20 litres per day with these supplies. For clarify this, we should have the panel data of weights and measurements of production of individual cows. Anyway, high dry period feeding for increasing the supplies of tissues for lactation period was not profitable (Agenäs 2003). Instead, high dry period feeding exposed cows with paralysis and ketosis (Roche 2007).

For silage comparisons, Khalili et al. (2005) explored spring harvesting to be significantly more productive than other harvests, while an analysis of these statistics defined that the second harvest of silage was the most productive. Furthermore, a threshold of 20 % in silage was used as a cut-off value to provide enough distinction between grass silage and clover silage. In the analysis, the clover silage received a coefficient of -0,019, representing 0,55 litres of milk per day per cow. However, there was a significant difference in the D-values of these silages, which may be due to the exceptional summers; exceptional rainy 2017 and exceptional dry 2018 (FMI 2018; FMI 2017). According to Juutinen (2011), a decrease in D-value from 690 to 650 reduced milk amount by 2,2 kg per day. In relation to this, the difference between the D-values for clover silage 670 and grass silage 685 affected 0.83 litres per day. In this case, clover silage would have a positive effect on milk amount. That was in line with the results of the study presented in this paper: the silage, of which 50 % consist of clover, increased energy-corrected milk amount by 2,3 kg compared to the ryegrass silage (Johansen et el. 2017).

7 Conclusions

The questions addressed in the thesis were: What are the responses to feed protein and energy and what are the differences in productivity and efficiency between the MoC:s? What background variables affect the differences in productivity and efficiency? How can the interdependence of background variables influence the conclusions?

Quantitative statistical analysis is an excellent tool for handling large statistical materials. However, it is a prerequisite that there is statistical material to deal with. Often this was the cause of the problems encountered in this thesis. Economic efficiency is impossible to determine if input prices or unit costs are not available.

The translog model used in this study yielded a protein response of 0,19, which represents a 0,19 % increase in milk yield as a percentage increase of the protein input. Similarly, an energy response of 0,94 means 0,94 % increase in output as a percentage of energy is added. The differences in efficacy between MoC:s are not very large, as evidenced by an average efficiency of 0,93.

The environmental factor z was used to test the effect of feeding, barn and milking type, concentrated feed sharing methods, silage harvest round and clover content on milk yield. The differences between the options were not so large that their output impact outweighed the expected cost effect. It is worth looking at the impact on returns when the investment is timely for other reasons. The effect of clover content on milk yield was slightly positive.

The effect of different environmental factors on the analytical results may not be observed separately if they have significant adverse effects and change at the same time. According to Juutinen (2011), the D-value influenced the silage production effect, which must be considered when studying the production response of clover and grass silage. Similarly, according to her, a decrease in the digestibility of silage by 10 g/kg reduced the daily protein amount by 18,2 g and the fat amount by 24,3 g. An increase the amount of concentrated feed per kilo of dry matter increased the protein amount by 30,3 g/day and the fat amount by 24,9 g/day. The average for the concentrated feed volume was 9,8 kg and standard deviation was 1,9 kg, and mean D-value was 682 with standard deviation of 31,5. In this case, the variation in the amount of concentration feed has similar effect on the analytical result for the protein amount as the variation of the D-value, the effect on the fat amount was smaller.

One of the contributions of this thesis was the emphasis on the technical efficiency in the evaluation of milk production performance. Section 2.2 outlines the current way of measuring the efficiency of milk production, which is mainly based on the measurement of milk volumes. Efficiency is thus based on the annual milk amount of cows or the annual energy-corrected milk amount of cows. The difference between the best and worst producer prices, based on producers’ milk quality pricing, was 39,2 %. The data was divided into three groups based on average milk amount, the low amount group below 25 litres, the middle group from 25 to 35 litres and the high milk amount group more than 35 litres per day per cow. The high-amount group had 95 MoC:s that achieved at least 0,96 effectiveness, the middle-amount group 279 calculations achieved the same efficiency, and low-amount group 16 calculations achieved that efficiency. In high-amount group, efficiency is likely to be achieved because of the amount of milk, but in the other two groups, efficiency cannot be justified based on milk amount.

Finally, the most significant contribution of the thesis is to apply protein and energy as inputs instead of different forages. And the following economic assessments. The price of protein would appear to have a significant impact on the dairy farm's financial performance. No farm-specific input prices for silage and concentrated feed were available in this study. In the future, special attention should be paid to the cost of protein production through silage harvesting, if we think of a cow as a ruminant for which silage is an essential part of the diet. Similarly, protein in concentrated feed usually faces a significantly longer chain of processes and transportations than in silage. This is likely to put more economic and environmental pressure on reducing concentrated feed content in diet in the future. The essential question will be: Is it more profitable to process feed outside or inside the cow?

References

Agenäs, S. 2003, Effects of Feeding Intensity During the Dry Period. 1. Feed Intake, Body Weight, and Milk Production, Journal of dairy science, 86(3), s. 870-882. DOI 10.3168/jds. S0022-0302(03)73670-4.

Aigner, D., Lovell, K., Schmidt, P. 1977, Formulation and estimation of stochastic frontier production function models, Journal of Econometrics, **6**(1), s. 21-37. DOI 10.1016/0304-4076(77)90052-5.

Alstrup, L., Søegaard, K. & Weisbjerg, M.R. 2016, Effects of maturity and harvest season of grass-clover silage and of forage-to-concentrate ratio on milk production of dairy cows, Journal of Dairy Science, 99(1), s. 328-340. DOI 10.3168/jds.2015-9802.

Aspila, P., Kaustell, K., Mälkiä, P., Rautala, H., Salmela, P.& Suvitie, M. 1994, Tieto tuottamaan no 70, Maaseutukeskusten liitto, Kokemäki.

Coelli, T., Rao, D.S.P., O´Donnell, C.J. & Battese, G.E. 2005, An introduction to efficiency and productivity analysis, Springer Science & Business Media, Inc., 233 Spring Street, New York, NY 10013, USA

Dawson, J. 2017, Analysing quantitative survey data for business and management students, SAGE Publications Ltd, London.

Dewhurst, R.J., Moorby, J.M., Dhanoa, M.S., Evans, R.T.& W. J. Fisher, W.J. 2000, Effects of Altering Energy and Protein Supply to Dairy Cows During the Dry Period. 1. Intake, Body Condition, and Milk Production, Journal of dairy science, s. 1782-1794, DOI 10.3168/jds.S0022-0302(00)75049-1.

Ettala, E., Lampila, M. 1978, Factors affecting voluntary silage intake by dairy cows. Annales Agriculturae Fenniae, **17**(4), s. 163-174.

Finnish Meteorological Institute, FMI, https://ilmatieteenlaitos.fi/vuosi-2018, https://ilmatieteenlaitos.fi/vuosi-2017, Read 14.3.2020

Grant, R.J. & Ferraretto, L.F. 2018, Silage review: Silage feeding management: Silage characteristics and dairy cow feeding behavior, Journal of Dairy Science, 101(5), s. 4111-4121. DOI 10.3168/jds.2017-13729.

Gravelle, H., Rees, R. 1992, Microeconomics, Pearson Education Limited, Edinburgh Gate, Harlow, Essex CM20 2JE, England. http://www.pearsoneduc.com.

Greene, W.H. 1980, On the estimation of a flexible frontier production model, Journal of Econometrics, **13**(1), s. 101-115. DOI 10.1016/0304-4076(80)90045-7.

Hardaker, J.B., Anderson, J.R. & Huirne, R.B.M. 1997, Coping with risk in agriculture, CAB International, Wallingford.

Henningsen, Arne (2019): Introduction to Econometric Production Analysis with R. Collection of Lecture Notes. 4th Draft Version. Department of Food and Resource Economics, University of Copenhagen. Available at http://leanpub.com/ProdEconR/.

Hirsjärvi, S., Remes, P. & Sajavaara, P. 1997, Tutki ja kirjoita, 1.-2. p. p., Kirjayhtymä, Helsinki.

Huang, C., Liu, J.-T., 1994, Estimation of a non-neutral stochastic frontier production function, Journal of Productivity Analysis, **5**(2), s. 171-180. DOI 10.1007/BF01073853.

Huhtanen, P. 1998, Supply of nutrients and productive responses in dairy cows given diets based on restrictively fermented silage, Agricultural and Food Science, **7**(2). DOI 10.23986/afsci.72861.

Huhtanen, P., Rinne, M. & Nousiainen, J. 2008, Evaluation of concentrate factors affecting silage intake of dairy cows: a development of the relative total diet intake index, animal, **2**(6), s. 942-953. DOI 10.1017/S1751731108001924.

The International Committee for Animal Recording (ICAR), www.icar.org. ICAR, Via Savoia 78, sc. A int. 3, 00198 Rome, Italy.

Johansen, M., Søegaard, K., Lund, P. & Weisbjerg, M.R. 2017, Digestibility and clover proportion determine milk production when silages of different grass and clover species are fed to dairy cows, Journal of Dairy Science, 100(11), s. 8861-8880. DOI 10.3168/jds.2017-13401.

Jondrow, J., Lovell, C. A. K., Materov, I. S., Scmidt, P. 1982, On the estimation of technical inefficiency in the stochastic frontier production function model, Journal of Econometrics, **19**(2), s. 233-238. DOI 10.1016/0304-4076(82)90004-5.

Juutinen, E. 2011, Säilörehun sulavuuden sekä väkirehu- ja valkuaistäydennyksen vaikutukset maidontuotannossa., Maisterin tutkielma, Helsingin yliopisto, Maataloustieteiden laitos, Helsinki.

Khalili, H., Sairanen, A., Nousiainen, J. & Huhtanen, P. 2005, Effects of silage made from primary or regrowth grass and protein supplementation on dairy cow performance, Livestock Production Science, 96(2), s. 269-278. DOI 10.1016/j.livprodsci.2005.02.007.

Kessler, E.C. 2014, Milk production during the colostral period is not related to the later lactational performance in dairy cows, Journal of dairy science, 97(4), s. 2186-2192. DOI 10.3168/jds.2013-7573.

Kokkonen, T., Tesfa, A., Tuori, M. & Syrjälä-Qvist, L. 2004, Concentrate feeding strategy of dairy cows during transition period, Livestock Production Science, 86(1), s. 239-251. DOI 10.1016/j.livprodsci.2003.09.003.

Lothgren, M. 2000, Specification and estimation of stochastic multiple-output production and technical inefficiency, Applied Economics, **32**(12), s. 1533-1540. DOI 10.1080/000368400418943.

Maaseudun Tulevaisuus, Markkinakatsaus, 6.4.2020. Read 10.4.2020

McLeod, M. N., Kennedy, P. M. & Minson, D. J. 1990, Resistance of leaf and stem fractions of tropical forage to chewing and passage in cattle, British Journal of Nutrition, Volume 63, Issue 1, s. 105-119. DOI: 10.1079/BJN19900096.

McNamara, S., Butler, S.T., Butler, T., Ryan, D.P., Mee, J.F., Dillon, P., O'Mara, F.P., Anglesey, D., Rath, M. & Murphy, J.J. 2003, Effect of offering rumen-protected fat supplements on fertility and performance in spring-calving Holstein–Friesian cows, Animal Reproduction Science, 79(1), s. 45-56. DOI 10.1016/S0378-4320(03)00111-8.

Meeusen, W. 1977, Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error, International Economic Review, **18**(2), s. 435-444. DOI 10.2307/2525757.

Mertens, D.R. 1997, Creating a system for meeting the fiber requirements of dairy cows, Journal of dairy science, **80**(7), s. 1463. DOI 10.3168/jds.S0022-0302(97)76075-2.

Natural Resources Institute Finland, LUKE, http://statdb.luke.fi/PXWeb/pxweb/fi/LUKE/LUKE__02%20Maatalous__04%20Tuotanto__12%20Kotielainten%20lukumaara/03_Lypsylehmien_lukumaara_karjakokoluokka.px/?rxid=e88f813d-2a68-42cb-9a17-6d2a3e53c76e. Read 12.1.2020.

Natural Resources Institute Finland, LUKE, 2015, Rehutaulukot ja ruokintasuositukset: märehtijät - siat - siipikarja - hevoset, Luonnonvarakeskus, Helsinki.

Olson, K.D. 2004, Farm management: principles and strategies, Iowa State Press, Ames, IA.

ProAgria Keskusten Liitto, Tuottopuntari 2018. www.webwisu.fi/tuottopehtoori. Read 27.12.2019.

ProAgria Keskusten Liitto, https://proagria.fi/sites/default/files/attachment/lypsykarjan_tuotosseurannan_ohjesaanto_2019_0.pdf Read 10.1.2020.

Roche, J.R. 2007, Milk production responses to pre- and postcalving dry matter intake in grazing dairy cows, Livestock Science, 110(1), s. 12-24. DOI 10.1016/j.livsci.2006.08.016.

Ryhänen, M., Sipiläinen, T. 2018 Maatalousyrityksen johtaminen ja toiminnan kehittäminen, Tuotannon suunnittelu strategisen johtamisen tukena, Tempest oy, Helsinki.

Ryhänen, M., Huhtanen, P., Jaakkola, S. & Ahvenjärvi, S. 1996, EU-jäsenyyden vaikutus maidontuotantoon, Ylätalo, M., Maatalousyritysten sopeutuminen EU:ssa vallitseviin hintasuhteisiin, University of Helsinki, Department of Economics and Management Publications 12: 75-118.

Vanholder, T. 2015, Risk factors for subclinical and clinical ketosis and association with production parameters in dairy cows in the Netherlands, Journal of dairy science, 98(2), s. 880-888. DOI 10.3168/jds.2014-8362.

Varian, H.R. 1999, Intermediate microeconomics: a modern approach, 5. ed. p., Norton, New York.

Vauhkonen, E. 2011, Säilörehun korjuuajan vaikutus maitotilan talouteen – lyhyen aikavälin näkökulma, Maisterin tutkielma, Helsingin yliopisto, Taloustieteen laitos.

Vilja-alan yhteistyöryhmä, 2013, https://www.vyr.fi/document/1/75/02f1e1b/oppaat_87d6615_Kaura_rehuna_FI_12s_120613.pdf

Wang, H. 2002, One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels, Journal of Productivity Analysis, **18**(2), s. 129-144. DOI 10.1023/A:1016565719882.

Appendix 1

In general, costs are considered at certain input price levels. In this paper, the calculated cost of concentrated feed changes linearly between the estimated wheat and rapeseed prices. The price depends on the required protein content of the concentrated feed, calculated according to the model in this appendix.

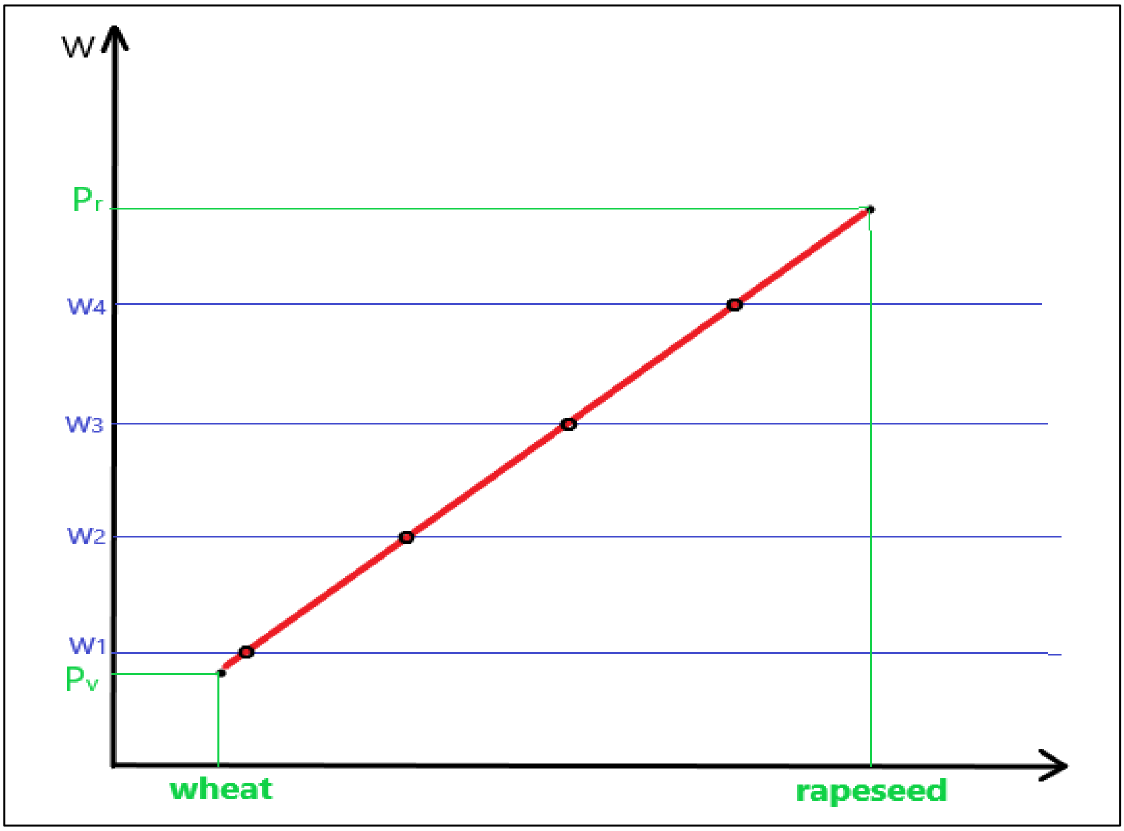


Figure 12. Prices of wheat and rapeseed compared to different cost levels.

Figure 12 illustrates the relationship between the fixed price of concentrated feed and the price of concentrated feed relative to wheat and rapeseed prices. In the Figure 7, w_1, w_2, w_3 and w_4 depict fixed price levels drawn in blue thin lines. The red solid line represents the price of concentrated feed relative to the price of wheat and rapeseed. The price for wheat is fixed at EUR 160/tonne and that of rapeseed at EUR 300/tonne. According to chapter 3.2, the basic nutrients of cows are energy and protein. In general, the most important factor affecting the price of concentrated feed is its protein content, which is regulated by the relationship between the grain and the source of protein. In this case, rapeseed is used as the protein source. Cereal starch and silage are generally enough sources of energy.

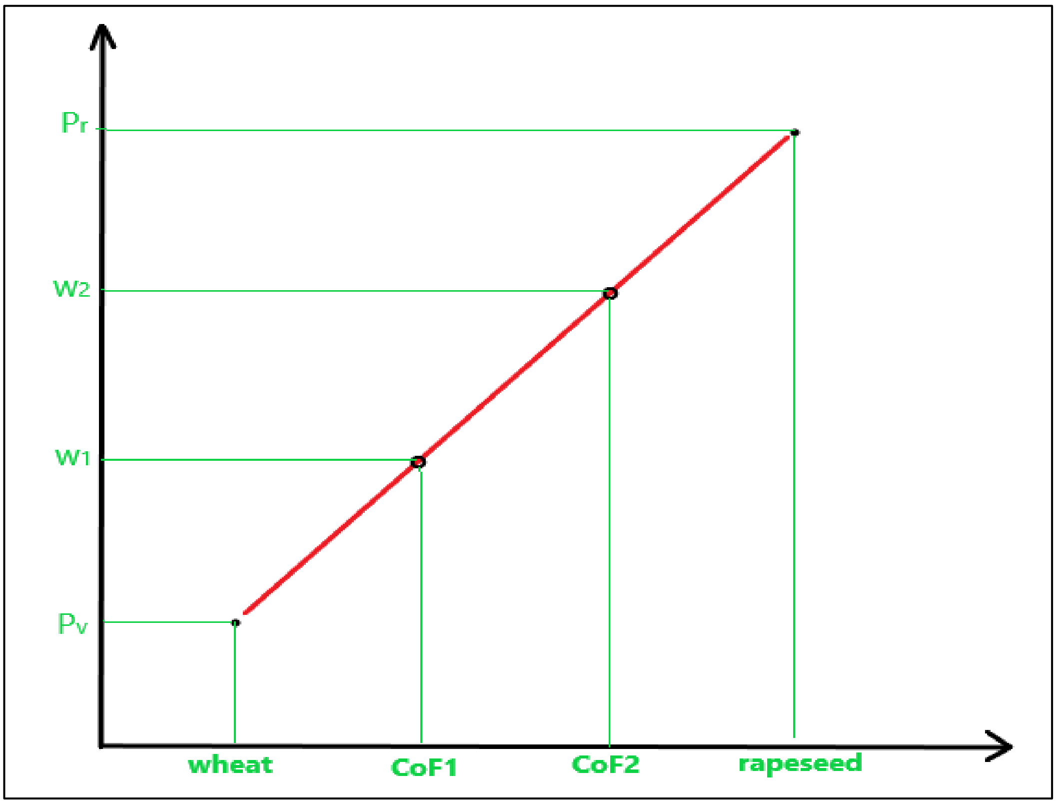


Figure 13. Costs of concentrated feeds compared to price of wheat and rapeseed.

In this study, the cost of concentrated feed is defined for each MoC based on the input data for that calculation. Figure 13 illustrates the effect of the relationship between wheat and rapeseed used in concentrated feed on the cost of concentrated feed. The y-axis represents the price of concentrated feed and the x-axis shows the transition of the concentrated feed ratio from wheat alone to rapeseed alone. Concentrated feed 1 (CoF1 in figure) has less rapeseed than concentrated feed 2 (CoF2 in figure) because of its lower protein content requirement, resulting a lower price w_1 for concentrated feed 1 than w_2 for concentrated feed 2. The required protein content of the concentrated feed is calculated as follows. First, we formulate the relationship between protein concentrations:

$$V_d = V_s * (1 - C_v) + V_c * C_v \qquad (1)$$

In equation V_d is the protein content of the whole diet, V_s is the protein content of the silage, V_c is the protein content of the concentrated feed and C_v is the percentage of the concentrated feed divided by 100. From the statistics used in the thesis, these starting values are used to calculate the protein concentration V_c of the concentrated feed solved by the equation:

$$V_c = \frac{V_d - V_s * (1 - C_v)}{C_v} \tag{2}$$

The change in protein content ΔV is calculated by subtracting the protein content V_c of the required concentrated feed from the protein content V_v of the wheat. Finally, the dry correction D_v must be made since the statistics indicate the quantities expressed as dry matter and marked prices are used as the basis for pricing, determined on average by 14 per cent humidity. At the same time the wheat price P_v is added at the starting level of EUR 169 per tonne:

$$w_v = \frac{((V_d - V_s * (1 - C_v))/C_v) - V_v) * (P_r - P_v)}{(V_r - V_v) * D_v} + P_v \tag{3}$$

The total calculatory cost of concentrated feed is determined by the amount of concentrated feed and the price as shown below:

$$W_v = \frac{A_v * w_v}{1000} \tag{4}$$

Cereal and rapeseed prices have fluctuated dramatically in recent years, so the calculation must be adjusted to the prevailing level. The purpose of the prices used in this paper is to illustrate the impact of the protein content of silage to the cost of concentrated feed.

Appendix 2

Table 4. Classification of statistical variables by different characteristics (Davson 2017).

Major type	Sub-type	Description	Examples
Categorical	Nominal	Used to distinguish between groups that have no natural order.	Type of milker, type of forage sharing, type of sharing concentrated feed. A categorizing variable is used (1, 2, 3, and so on).
	Binary	Special case of nominal rating; there are only two classes.	Type of barn, type of silage. Dummy variable is used (0 tai 1).
	Ordinal	Used to distinguish between groups that have a natural order. Whenever possible it is possible to rank the material in order of merit.	Groups (quantile) arranged; example best quantile (Q_{1}, Q_{2}, ..). or individually with range (1., 2.,3., ..).
Numerical	Continuous, interval	Used to express a numerical quantity, a graduated quantity. (Integer, usually not zero)	Milk amount of dairy farm, D-value
	Continuous, ratio	Used to express a continuous amount, any positive value can be obtained. Can also get a value of zero	Protein and fat percentage.
	Discrete	Used as a unit of account when the population is small. Only integers	Cows per herd

Appendix 3

Analyses with R-studio:

CD-function

```
final maximum likelihood estimates
                       Estimate  Std. Error  z value  Pr(>|z|)    
(Intercept)         -2.95259491  0.18079670 -16.3310 < 2.2e-16 ***
log(Wholeprotein)    0.20294275  0.02687221   7.5521 4.282e-14 ***
log(Energyamount)    0.94887403  0.03814262  24.8770 < 2.2e-16 ***
sigmaSq              0.01412202  0.00077046  18.3293 < 2.2e-16 ***
gamma                0.73178867  0.03053407  23.9663 < 2.2e-16 ***
sigmaSqU             0.01033434  0.00094499  10.9359 < 2.2e-16 ***
sigmaSqV             0.00378769  0.00029325  12.9161 < 2.2e-16 ***
sigma                0.11883611  0.00324169  36.6587 < 2.2e-16 ***
sigmaU               0.10165794  0.00464789  21.8719 < 2.2e-16 ***
sigmaV               0.06154418  0.00238246  25.8322 < 2.2e-16 ***
lambdaSq             2.72840325  0.42445385   6.4280 1.293e-10 ***
lambda               1.65178790  0.12848316  12.8561 < 2.2e-16 ***
varU                 0.00375529          NA       NA        NA    
sdU                  0.06128045          NA       NA        NA    
gammaVar             0.49785277          NA       NA        NA    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2010.65 

cross-sectional data
total number of observations = 1943 

mean efficiency: 0.9245748
```

CD-function with confeedwholecost

```
final maximum likelihood estimates
                           Estimate  Std. Error  z value  Pr(>|z|)    
(Intercept)             -3.04371424  0.18368832 -16.5700 < 2.2e-16 ***
log(Wholeprotein)        0.22672474  0.02858838   7.9307 2.180e-15 ***
log(Energyamount)        0.96274863  0.03840312  25.0695 < 2.2e-16 ***
log(Confeedwholecost) -0.02377686  0.00948321  -2.5073   0.01217 *  
sigmaSq                  0.01428325  0.00076387  18.6986 < 2.2e-16 ***
gamma                    0.74230624  0.02878440  25.7885 < 2.2e-16 ***
sigmaSqU                 0.01060255  0.00092779  11.4277 < 2.2e-16 ***
sigmaSqV                 0.00368070  0.00028150  13.0751 < 2.2e-16 ***
sigma                    0.11951256  0.00319577  37.3971 < 2.2e-16 ***
sigmaU                   0.10296867  0.00450523  22.8554 < 2.2e-16 ***
sigmaV                   0.06066881  0.00232001  26.1502 < 2.2e-16 ***
lambdaSq                 2.88057509  0.43346033   6.6455 3.021e-11 ***
lambda                   1.69722570  0.12769673  13.2911 < 2.2e-16 ***
varU                     0.00385276          NA       NA        NA    
sdU                      0.06207057          NA       NA        NA    
gammaVar                 0.51141912          NA       NA        NA    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2013.839 

cross-sectional data
total number of observations = 1943 

mean efficiency: 0.9237071
```

```
Likelihood ratio test

Model 1: prodCDSfa
Model 2: prodCDSfac
  #Df LogLik Df  Chisq Pr(>Chisq)
1   5 2010.7
2   6 2013.8  1 6.3778    0.01156 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

CD-function with truncNorm

```
final maximum likelihood estimates
                         Estimate  Std. Error  z value  Pr(>|z|)
(Intercept)           -3.17450094  0.18209063 -17.4336 < 2.2e-16 ***
log(Wholeprotein)      0.22935788  0.02854696   8.0344 9.403e-16 ***
log(Energyamount)      0.98271636  0.03797012  25.8813 < 2.2e-16 ***
log(Confeedwholecost) -0.02526408  0.00944423  -2.6751  0.007471 **
sigmaSq                0.03316823  0.00372387   8.9069 < 2.2e-16 ***
gamma                  0.87546021  0.01876121  46.6633 < 2.2e-16 ***
mu                    -0.34080765  0.06840898  -4.9819 6.296e-07 ***
sigmaSqU               0.02903746  0.00379461   7.6523 1.974e-14 ***
sigmaSqV               0.00413076  0.00034781  11.8766 < 2.2e-16 ***
sigma                  0.18212146  0.01022358  17.8139 < 2.2e-16 ***
sigmaU                 0.17040382  0.01113417  15.3046 < 2.2e-16 ***
sigmaV                 0.06427102  0.00270578  23.7532 < 2.2e-16 ***
lambdaSq               7.02956227  1.20960759   5.8114 6.194e-09 ***
lambda                 2.65133217  0.22811317  11.6229 < 2.2e-16 ***
varU                   0.00331838          NA       NA        NA
sdU                    0.05760534          NA       NA        NA
gammaVar               0.44547096          NA       NA        NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2030.513

cross-sectional data
total number of observations = 1943

mean efficiency: 0.9405942
```

Translog function without confeedwholecost

```
final maximum likelihood estimates
                                        Estimate  Std. Error z value  Pr(>|z|)
(Intercept)                          -3.3325e+01 1.1310e+01 -2.9466  0.003213 **
log(Wholeprotein)                     1.0204e+00 2.8688e+00  0.3557  0.722072
log(Energyamount)                     1.1907e+01 4.7510e+00  2.5061  0.012206 *
I(0.5 * (log(Wholeprotein))^2)        4.8286e-01 4.5109e-01  1.0704  0.284429
I(0.5 * (log(Energyamount))^2)       -1.9500e+00 1.0018e+00 -1.9464  0.051601 .
I(log(Wholeprotein) * log(Energyamount)) -2.6259e-01 6.1869e-01 -0.4244  0.671253
sigmaSq                               1.2329e-02 8.0960e-04 15.2283 < 2.2e-16 ***
gamma                                 6.4109e-01 4.8422e-02 13.2397 < 2.2e-16 ***
sigmaSqU                              7.9040e-03 1.0797e-03  7.3203 2.475e-13 ***
sigmaSqV                              4.4249e-03 3.7215e-04 11.8903 < 2.2e-16 ***
sigma                                 1.1104e-01 3.6457e-03 30.4566 < 2.2e-16 ***
sigmaU                                8.8904e-02 6.0725e-03 14.6405 < 2.2e-16 ***
sigmaV                                6.6520e-02 2.7972e-03 23.7806 < 2.2e-16 ***
lambdaSq                              1.7862e+00 3.7591e-01  4.7518 2.016e-06 ***
lambda                                1.3365e+00 1.4063e-01  9.5036 < 2.2e-16 ***
varU                                  2.8721e-03         NA      NA        NA
sdU                                   5.3592e-02         NA      NA        NA
gammaVar                              3.9360e-01         NA      NA        NA
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2031.829

cross-sectional data
total number of observations = 1943

mean efficiency: 0.9332523
```

Likelihood ratio test

```
Model 1: prodTLSfah
Model 2: prodCDSfac
  #Df LogLik Df Chisq Pr(>Chisq)
1   8 2031.8
2   6 2013.8 -2 35.98  1.538e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Translog function with confeedwholecost

```
final maximum likelihood estimates
                                             Estimate  Std. Error z value  Pr(>|z|)
(Intercept)                               -4.2994e+01  1.1472e+01 -3.7476 0.0001785 ***
log(Wholeprotein)                          2.1282e+00  3.0008e+00  0.7092 0.4781903
log(Energyamount)                          1.5404e+01  4.7653e+00  3.2325 0.0012272 **
log(Confeedwholecost)                     -2.1339e+00  9.4375e-01 -2.2611 0.0237519 *
I(0.5 * (log(Wholeprotein))^2)            -1.0405e+00  5.6136e-01 -1.8536 0.0637984 .
I(0.5 * (log(Energyamount))^2)            -2.6495e+00  9.9643e-01 -2.6590 0.0078381 **
I(0.5 * (log(Confeedwholecost))^2)         3.5196e-02  1.9350e-02  1.8189 0.0689236 .
I(log(Wholeprotein) * log(Energyamount))  -1.4359e-01  6.5552e-01 -0.2191 0.8266104
I(log(Confeedwholecost) * log(Energyamount)) 3.1417e-01  1.9455e-01  1.6148 0.1063507
I(log(Wholeprotein) * log(Confeedwholecost)) 3.1826e-01  1.0395e-01  3.0618 0.0021999 **
sigmaSq                                    1.1724e-02  8.0305e-04 14.5993 < 2.2e-16 ***
gamma                                      6.3683e-01  5.1210e-02 12.4357 < 2.2e-16 ***
sigmaSqU                                   7.4661e-03  1.0789e-03  6.9201 4.513e-12 ***
sigmaSqV                                   4.2578e-03  3.6932e-04 11.5288 < 2.2e-16 ***
sigma                                      1.0828e-01  3.7083e-03 29.1986 < 2.2e-16 ***
sigmaU                                     8.6407e-02  6.2431e-03 13.8402 < 2.2e-16 ***
sigmaV                                     6.5252e-02  2.8299e-03 23.0576 < 2.2e-16 ***
lambdaSq                                   1.7535e+00  3.8826e-01  4.5163 6.293e-06 ***
lambda                                     1.3242e+00  1.4660e-01  9.0326 < 2.2e-16 ***
varU                                       2.7130e-03          NA      NA        NA
sdU                                        5.2087e-02          NA      NA        NA
gammaVar                                   3.8920e-01          NA      NA        NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2075.989

cross-sectional data
total number of observations = 1943

mean efficiency: 0.9349945
```

Likelihood ratio test

```
Model 1: prodCDSfac
Model 2: prodTLSfahe
  #Df LogLik Df Chisq Pr(>Chisq)
1   6 2013.8
2  12 2076.0  6 124.3  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Translog function with confeedwholecost and truncNorm

```
final maximum likelihood estimates
                                             Estimate  Std. Error  z value  Pr(>|z|)
(Intercept)                               -5.2064e+01 2.4224e+00 -21.4930 < 2.2e-16 ***
log(wholeprotein)                         -2.1718e-03 1.7970e+00  -0.0012 0.9990357
log(Energyamount)                          1.9210e+01 1.1090e+00  17.3217 < 2.2e-16 ***
log(Confeedwholecost)                     -2.0616e+00 8.7969e-01  -2.3435 0.0191015 *
I(0.5 * (log(wholeprotein))^2)            -1.2527e+00 4.6416e-01  -2.6988 0.0069585 **
I(0.5 * (log(Energyamount))^2)            -3.4469e+00 2.6977e-01 -12.7775 < 2.2e-16 ***
I(0.5 * (log(Confeedwholecost))^2)         3.9720e-02 1.7691e-02   2.2452 0.0247572 *
I(log(wholeprotein) * log(Energyamount))   2.9489e-01 4.1512e-01   0.7104 0.4774748
I(log(Confeedwholecost) * log(Energyamount)) 2.9676e-01 1.8145e-01 1.6355 0.1019380
I(log(wholeprotein) * log(Confeedwholecost)) 3.3440e-01 1.0019e-01 3.3376 0.0008451 ***
sigmaSq                                    2.6913e-02 3.7114e-03   7.2513 4.128e-13 ***
gamma                                      8.3938e-01 2.8199e-02  29.7657 < 2.2e-16 ***
mu                                        -3.0060e-01 6.8185e-02  -4.4086 1.040e-05 ***
sigmaSqU                                   2.2590e-02 3.8151e-03   5.9212 3.196e-09 ***
sigmaSqV                                   4.3228e-03 3.3688e-04  12.8318 < 2.2e-16 ***
sigma                                      1.6405e-01 1.1312e-02  14.5026 < 2.2e-16 ***
sigmaU                                     1.5030e-01 1.2692e-02  11.8424 < 2.2e-16 ***
sigmaV                                     6.5748e-02 2.5619e-03  25.6635 < 2.2e-16 ***
lambdaSq                                   5.2257e+00 1.0930e+00   4.7811 1.744e-06 ***
lambda                                     2.2860e+00 2.3907e-01   9.5621 < 2.2e-16 ***
varU                                       2.5815e-03         NA       NA        NA
sdU                                        5.0809e-02         NA       NA        NA
gammaVar                                   3.7390e-01         NA       NA        NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2084.733

cross-sectional data
total number of observations = 1943

mean efficiency: 0.947105
```

Likelihood ratio test

```
Model 1: prodCDSfactn
Model 2: prodTLSfahetn
  #Df LogLik Df  Chisq Pr(>Chisq)
1   7 2030.5
2  13 2084.7  6 108.44  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Likelihood ratio test

```
Model 1: OLS (no inefficiency)
Model 2: Error Components Frontier (ECF)
  #Df LogLik Df  Chisq Pr(>Chisq)
1  11 2063.7
2  13 2084.7  2 42.148  2.183e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Likelihood ratio test

```
Model 1: prodTLSfahe
Model 2: prodTLSfahetn
  #Df LogLik Df  Chisq Pr(>Chisq)
1  12 2076.0
2  13 2084.7  1 17.489   2.89e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Mean scaled translog function with confeedwholecost

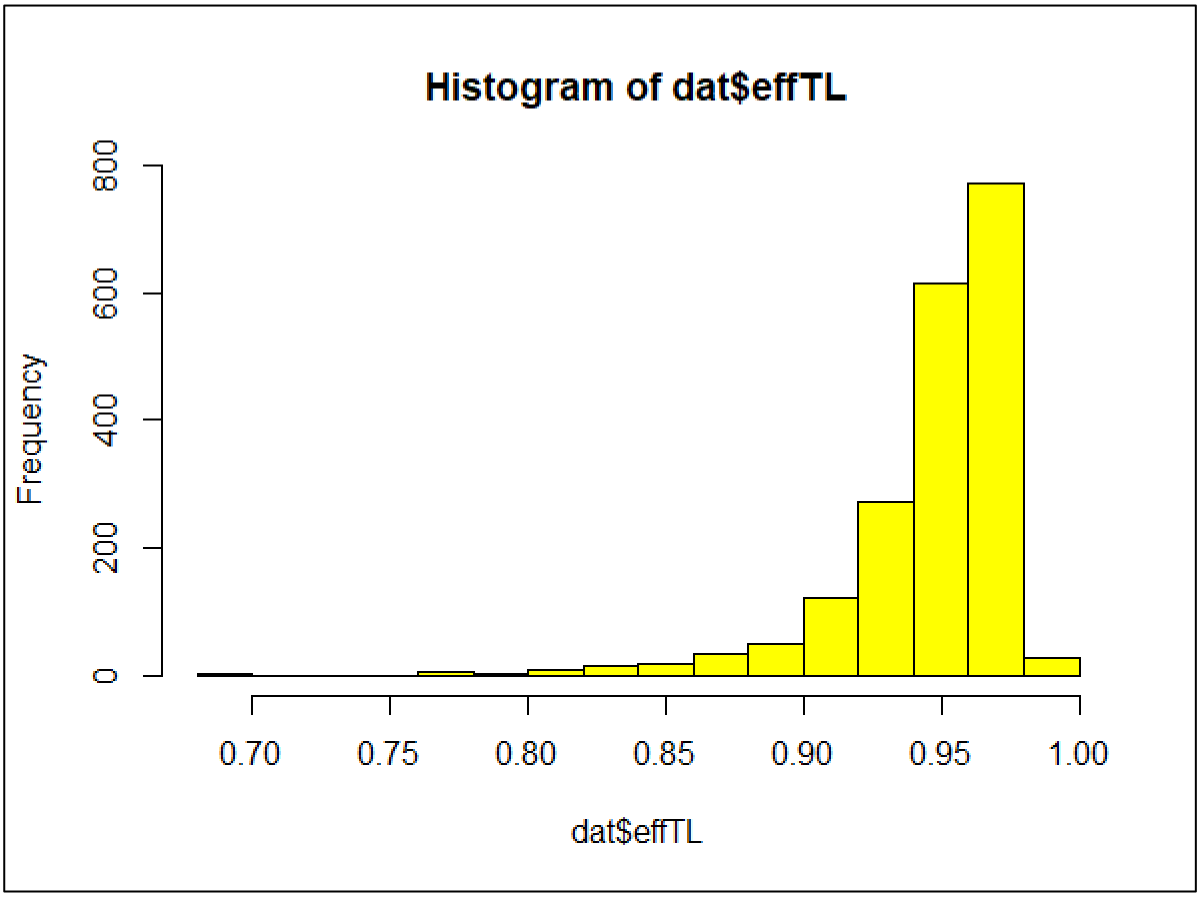
```
final maximum likelihood estimates
                             Estimate  Std. Error z value  Pr(>|z|)
(Intercept)                 0.07394612  0.00483852 15.2828 < 2.2e-16 ***
log(qmPro)                  0.19424233  0.03084776  6.2968 3.038e-10 ***
log(qmEne)                  0.93513264  0.03850738 24.2845 < 2.2e-16 ***
log(qmCoc)                  0.01817835  0.01260934  1.4417  0.149399
I(0.5 * log(qmPro)^2)      -1.04059353  0.55459792 -1.8763  0.060614 .
I(0.5 * log(qmEne)^2)      -2.64950398  0.98962598 -2.6773  0.007422 **
I(0.5 * log(qmCoc)^2)       0.03519678  0.01926931  1.8266  0.067764 .
I(log(qmPro) * log(qmEne)) -0.14351764  0.64672038 -0.2219  0.824379
I(log(qmPro) * log(qmCoc))  0.31827695  0.10276566  3.0971  0.001954 **
I(log(qmEne) * log(qmCoc))  0.31414885  0.19264922  1.6307  0.102958
sigmaSq                     0.01172407  0.00077722 15.0847 < 2.2e-16 ***
gamma                       0.63683058  0.05082057 12.5310 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 2075.989

cross-sectional data
total number of observations = 1943

mean efficiency: 0.934994
```

```
> all.equal(coef(prodTLmSfa)[-c(1:4)],coef(prodTLSfahe)[-c(1:4)], check.attribut
es = FALSE)
[1] "Mean relative difference: 0.001632445"
> all.equal(efficiencies(prodTLmSfa), efficiencies(prodTLSfahe))
[1] "Mean relative difference: 5.098843e-07"
```

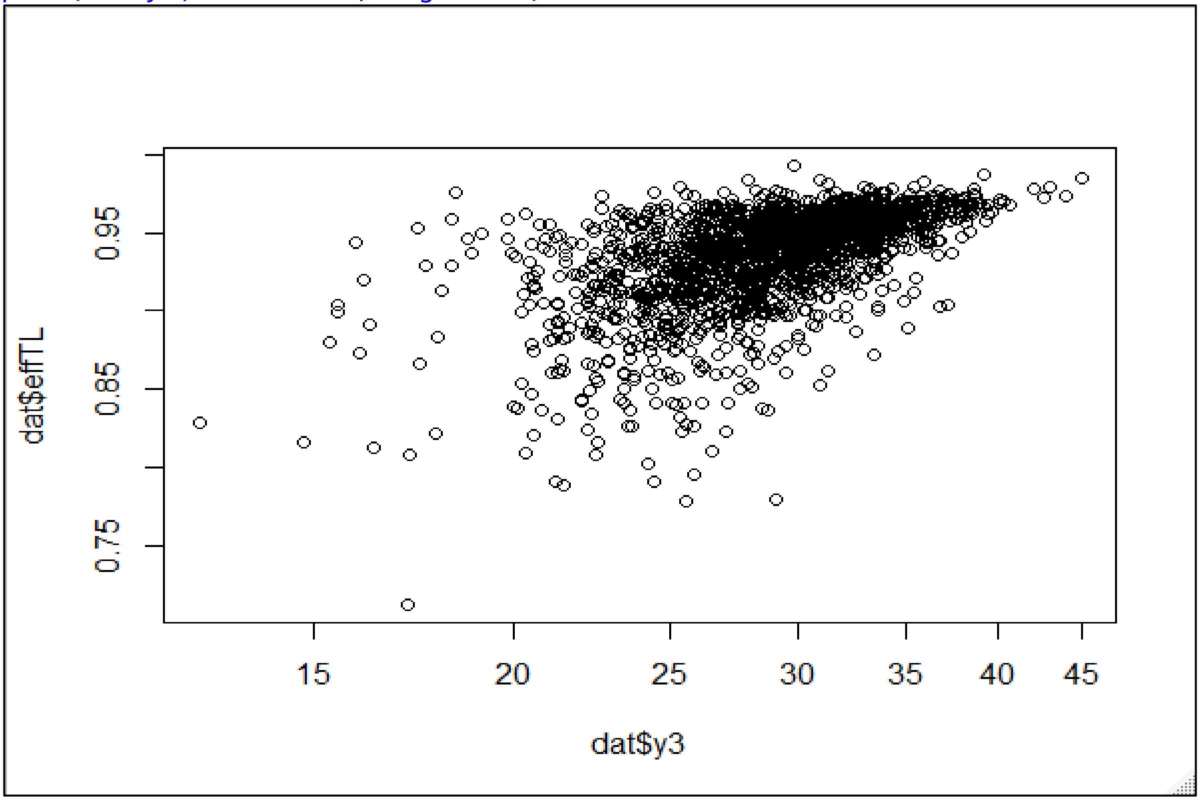
Mean-scaled translog function without confeedwholecost

```
final maximum likelihood estimates
                             Estimate  Std. Error z value  Pr(>|z|)
(Intercept)                 0.07332649  0.00469035 15.6335 < 2.2e-16 ***
log(qmPro)                  0.21161960  0.02910557  7.2708 3.575e-13 ***
log(qmEne)                  0.90833420  0.03905462 23.2580 < 2.2e-16 ***
I(0.5 * log(qmPro)^2)       0.48302800  0.44654986  1.0817   0.27939
I(0.5 * log(qmEne)^2)      -1.94962161  0.99225343 -1.9648   0.04943 *
I(log(qmPro) * log(qmEne)) -0.26286983  0.61162405 -0.4298   0.66735
sigmaSq                     0.01232881  0.00081233 15.1770 < 2.2e-16 ***
gamma                       0.64108062  0.04809440 13.3296 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
log likelihood value: 2031.829

cross-sectional data
total number of observations = 1943

mean efficiency: 0.9332534
```

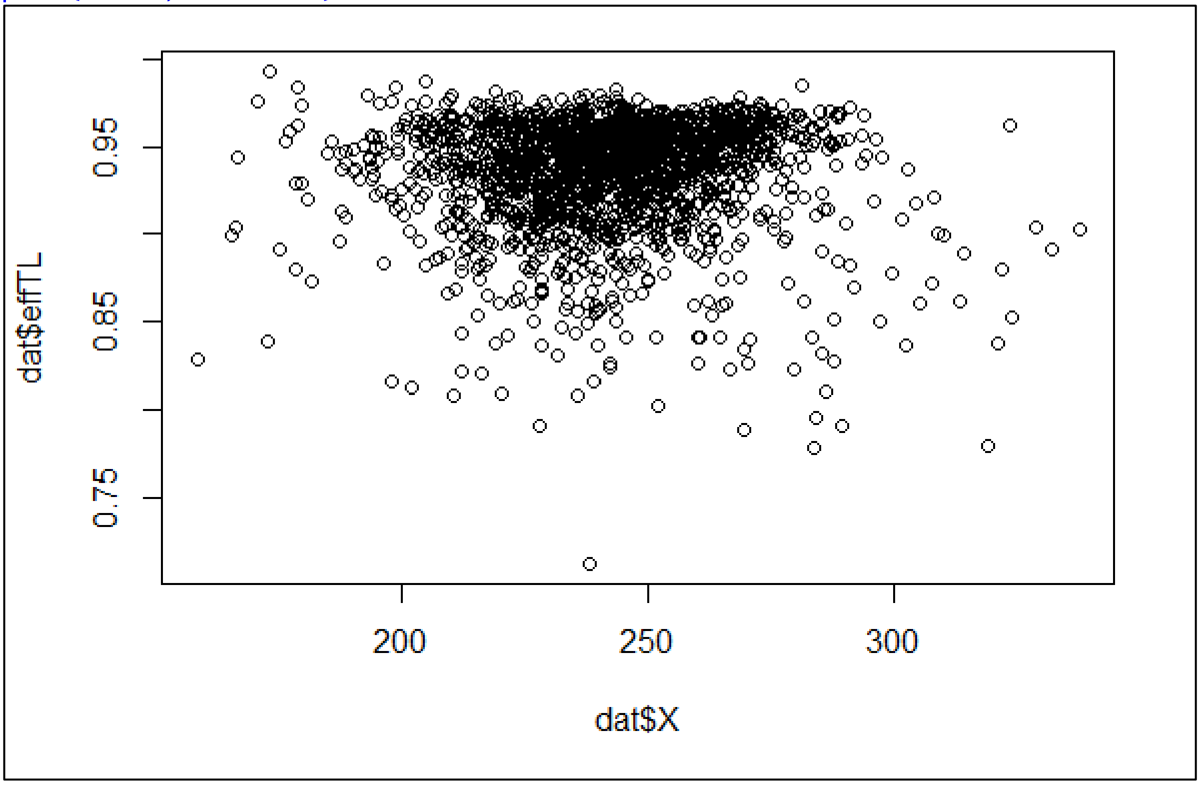
```
> all.equal(coef(prodTLmSfa)[-c(1:3)],coef(prodTLSfah)[-c(1:3)], check.attribute
s = FALSE)
[1] "Mean relative difference: 0.0002603897"
> all.equal(efficiencies(prodTLmSfa), efficiencies(prodTLSfah))
[1] "Mean relative difference: 1.181427e-06"
```



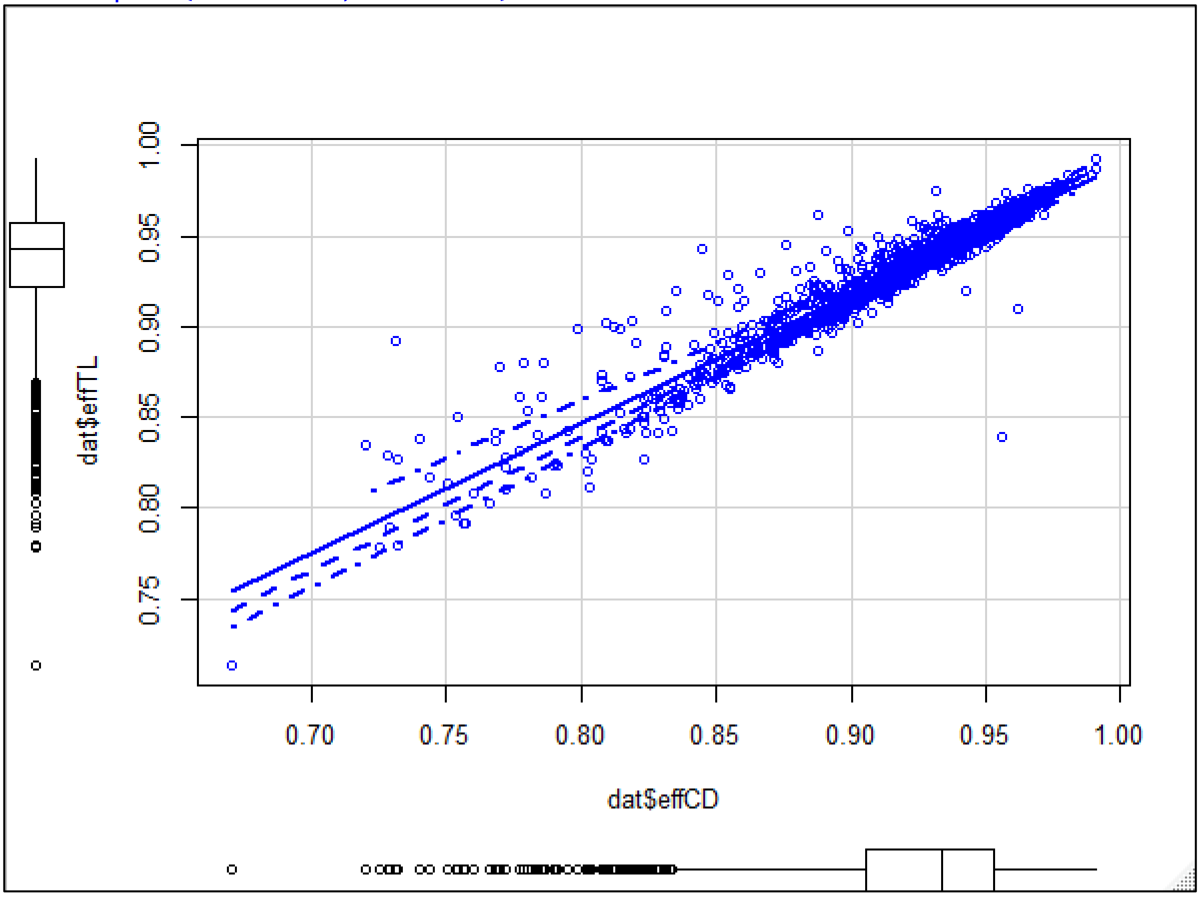

```
plot(dat$y3, dat$effTL, log = "x")
```



```
dat$X<-dat$Wholeprotein+dat$Energyamount+dat$Confeedwholecost
plot(dat$X,dat$effTL)
```




```
scatterplot(dat$effCD,dat$effTL)
```



Type of feeding

```
final maximum likelihood estimates
                                          Estimate  Std. Error z value  Pr(>|z|)    
(Intercept)                             -4.1667e+01 1.1491e+01 -3.6259 0.0002879 ***
log(wholeprotein)                        2.3339e+00 2.9875e+00  0.7812 0.4346688    
log(Energyamount)                        1.4868e+01 4.7698e+00  3.1171 0.0018263 ** 
log(Confeedwholecost)                   -2.1197e+00 9.4197e-01 -2.2503 0.0244298 *  
I(0.5 * (log(wholeprotein))^2)          -1.0215e+00 5.6163e-01 -1.8188 0.0689441 .  
I(0.5 * (log(Energyamount))^2)          -2.5411e+00 9.9680e-01 -2.5493 0.0107950 *  
I(0.5 * (log(Confeedwholecost))^2)       3.5408e-02 1.9359e-02  1.8290 0.0674046 .  
I(log(wholeprotein) * log(Energyamount)) -1.8627e-01 6.5330e-01 -0.2851 0.7755539    
I(log(Confeedwholecost) * log(Energyamount))  3.1180e-01 1.9422e-01  1.6054 0.1084079    
I(log(wholeprotein) * log(Confeedwholecost))  3.1713e-01 1.0401e-01  3.0491 0.0022953 ** 
second                                  -1.2319e-04 5.8365e-03 -0.0211 0.9831600    
third                                    8.1989e-03 5.1522e-03  1.5913 0.1115334    
sigmaSq                                  1.1764e-02 7.9943e-04 14.7153 < 2.2e-16 ***
gamma                                    6.4098e-01 5.0494e-02 12.6940 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2077.326

cross-sectional data
total number of observations = 1943

mean efficiency: 0.9347041
```

Separated feeding as reference

Mixed feeding coefficient -0,000123
Mixed feeding supplemented with concentrated feed 0,00820

Clover silage

```
final maximum likelihood estimates
                                          Estimate  Std. Error z value  Pr(>|z|)    
(Intercept)                             -4.1873e+01 1.1339e+01 -3.6927 0.0002219 ***
log(wholeprotein)                        2.7456e+00 2.9691e+00  0.9247 0.3551107    
log(Energyamount)                        1.4868e+01 4.7066e+00  3.1590 0.0015829 ** 
log(Confeedwholecost)                   -2.3172e+00 9.4388e-01 -2.4550 0.0140882 *  
I(0.5 * (log(wholeprotein))^2)          -9.2054e-01 5.5944e-01 -1.6455 0.0998746 .  
I(0.5 * (log(Energyamount))^2)          -2.5229e+00 9.8372e-01 -2.5646 0.0103298 *  
I(0.5 * (log(Confeedwholecost))^2)       2.9712e-02 1.9393e-02  1.5321 0.1255034    
I(log(wholeprotein) * log(Energyamount)) -2.8242e-01 6.4942e-01 -0.4349 0.6636483    
I(log(Confeedwholecost) * log(Energyamount))  3.5256e-01 1.9459e-01  1.8118 0.0700247 .  
I(log(wholeprotein) * log(Confeedwholecost))  2.9774e-01 1.0381e-01  2.8681 0.0041300 ** 
legume                                  -1.7662e-02 4.9436e-03 -3.5726 0.0003535 ***
sigmaSq                                  1.1636e-02 8.0068e-04 14.5324 < 2.2e-16 ***
gamma                                    6.3593e-01 5.1523e-02 12.3428 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2082.365

cross-sectional data
total number of observations = 1943

mean efficiency: 0.9352664
```

Grass silage as reference

Clover silage coefficient -0,0177

Silage harvest round

```
final maximum likelihood estimates
                                           Estimate  Std. Error z value  Pr(>|z|)    
(Intercept)                             -4.4373e+01  1.1357e+01 -3.9071 9.342e-05 ***
log(Wholeprotein)                        1.6958e+00  2.9764e+00  0.5697 0.5688562    
log(Energyamount)                        1.5972e+01  4.7172e+00  3.3860 0.0007093 ***
log(Confeedwholecost)                   -2.0952e+00  9.3982e-01 -2.2294 0.0257879 *  
I(0.5 * (log(Wholeprotein))^2)          -1.1026e+00  5.6006e-01 -1.9687 0.0489866 *  
I(0.5 * (log(Energyamount))^2)          -2.7672e+00  9.8640e-01 -2.8054 0.0050256 ** 
I(0.5 * (log(Confeedwholecost))^2)       3.5554e-02  1.9341e-02  1.8383 0.0660168 .  
I(log(Wholeprotein) * log(Energyamount)) -5.3000e-02  6.5084e-01 -0.0814 0.9350971    
I(log(Confeedwholecost) * log(Energyamount))  3.0597e-01  1.9378e-01  1.5789 0.1143531    
I(log(Wholeprotein) * log(Confeedwholecost))  3.2074e-01  1.0373e-01  3.0920 0.0019879 ** 
Summer                                   1.2869e-02  5.0548e-03  2.5459 0.0109003 *  
Autumn                                   1.0722e-02  5.8467e-03  1.8339 0.0666652 .  
sigmaSq                                  1.1729e-02  7.8051e-04 15.0269 < 2.2e-16 ***
gamma                                    6.4082e-01  4.9640e-02 12.9094 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2080.12 

cross-sectional data
total number of observations = 1943 

mean efficiency: 0.9348103
```

Spring harvest as a reference

Summer harvest coefficient 0,0129

Autumn harvest coefficient 0,0107

Type of barn

```
final maximum likelihood estimates
                                           Estimate  Std. Error z value  Pr(>|z|)    
(Intercept)                             -4.2287e+01  1.1448e+01 -3.6937  0.000221 ***
log(Wholeprotein)                        2.2236e+00  2.9780e+00  0.7467  0.455260    
log(Energyamount)                        1.5120e+01  4.7521e+00  3.1817  0.001464 ** 
log(Confeedwholecost)                   -2.1029e+00  9.4131e-01 -2.2340  0.025485 *  
I(0.5 * (log(Wholeprotein))^2)          -1.0170e+00  5.5981e-01 -1.8167  0.069257 .  
I(0.5 * (log(Energyamount))^2)          -2.5917e+00  9.9309e-01 -2.6097  0.009061 ** 
I(0.5 * (log(Confeedwholecost))^2)       3.5059e-02  1.9318e-02  1.8148  0.069551 .  
I(log(Wholeprotein) * log(Energyamount)) -1.6673e-01  6.5110e-01 -0.2561  0.797894    
I(log(Confeedwholecost) * log(Energyamount))  3.0869e-01  1.9404e-01  1.5909  0.111637    
I(log(Wholeprotein) * log(Confeedwholecost))  3.1656e-01  1.0367e-01  3.0535  0.002262 ** 
pihatto1                                 5.5264e-03  3.8488e-03  1.4359  0.151043    
sigmaSq                                  1.1793e-02  7.8259e-04 15.0686 < 2.2e-16 ***
gamma                                    6.4280e-01  4.9614e-02 12.9561 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2077.029 

cross-sectional data
total number of observations = 1943 

mean efficiency: 0.9345371
```

Tied-housed system as a reference

Loose-housed system coefficient 0,00553

Type of milker

```
final maximum likelihood estimates
                                             Estimate  Std. Error z value  Pr(>|z|)    
(Intercept)                               -4.1149e+01  1.1447e+01 -3.5948 0.0003246 ***
log(Wholeprotein)                          3.0466e+00  2.9965e+00  1.0167 0.3092799    
log(Energyamount)                          1.4540e+01  4.7546e+00  3.0581 0.0022275 ** 
log(Confeedwholecost)                     -2.3302e+00  9.4356e-01 -2.4696 0.0135265 *  
I(0.5 * (log(Wholeprotein))^2)            -8.7813e-01  5.6204e-01 -1.5624 0.1181967    
I(0.5 * (log(Energyamount))^2)            -2.4481e+00  9.9436e-01 -2.4620 0.0138159 *  
I(0.5 * (log(Confeedwholecost))^2)         3.3037e-02  1.9339e-02  1.7083 0.0875870 .  
I(log(Wholeprotein) * log(Energyamount))  -3.4892e-01  6.5517e-01 -0.5326 0.5943352    
I(log(Confeedwholecost) * log(Energyamount))  3.5552e-01  1.9454e-01  1.8275 0.0676259 .  
I(log(Wholeprotein) * log(Confeedwholecost))  2.9811e-01  1.0387e-01  2.8700 0.0041042 ** 
Asema1                                    -1.2966e-02  5.1102e-03 -2.5373 0.0111699 *  
Putki1                                    -9.0895e-03  4.5745e-03 -1.9870 0.0469235 *  
Auas1                                     -5.3073e-02  2.8161e-02 -1.8846 0.0594781 .  
sigmaSq                                    1.1538e-02  7.9311e-04 14.5475 < 2.2e-16 ***
gamma                                      6.2715e-01  5.2983e-02 11.8369 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2080.793 

cross-sectional data
total number of observations = 1943 

mean efficiency: 0.9359374
```

AMS as a reference

Milking parlor coefficient -0,0130
Pipeline milking coefficient -0,0091

AMS and milking parlor coefficient -0,0531

Concentrated feed sharing type

```
final maximum likelihood estimates
                                             Estimate  Std. Error z value  Pr(>|z|)    
(Intercept)                               -4.2436e+01  1.1472e+01 -3.6992 0.0002163 ***
log(Wholeprotein)                          2.1590e+00  2.9894e+00  0.7222 0.4701732    
log(Energyamount)                          1.5186e+01  4.7623e+00  3.1888 0.0014284 ** 
log(Confeedwholecost)                     -2.0843e+00  9.4461e-01 -2.2065 0.0273509 *  
I(0.5 * (log(Wholeprotein))^2)            -1.0730e+00  5.6217e-01 -1.9087 0.0562952 .  
I(0.5 * (log(Energyamount))^2)            -2.6091e+00  9.9541e-01 -2.6211 0.0087651 ** 
I(0.5 * (log(Confeedwholecost))^2)         3.6996e-02  1.9324e-02  1.9146 0.0555488 .  
I(log(Wholeprotein) * log(Energyamount))  -1.4273e-01  6.5374e-01 -0.2183 0.8271668    
I(log(Confeedwholecost) * log(Energyamount))  3.0303e-01  1.9474e-01  1.5561 0.1196837    
I(log(Wholeprotein) * log(Confeedwholecost))  3.2754e-01  1.0416e-01  3.1448 0.0016622 ** 
lypsrobki1                                 9.3062e-03  1.0037e-02  0.9271 0.3538511    
lypsrob1                                   1.2978e-02  7.1496e-03  1.8152 0.0694917 .  
kioski1                                    7.6732e-03  6.4113e-03  1.1968 0.2313776    
kisko1                                     4.1573e-03  6.1865e-03  0.6720 0.5015848    
hand1                                      9.1668e-03  7.0405e-03  1.3020 0.1929150    
sigmaSq                                    1.1666e-02  8.0391e-04 14.5121 < 2.2e-16 ***
gamma                                      6.3451e-01  5.1883e-02 12.2295 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
log likelihood value: 2078.137 

cross-sectional data
total number of observations = 1943 

mean efficiency: 0.9352646
```

Mixed feed as a reference

Parlor/AMS +kiosk coefficient 0,0093

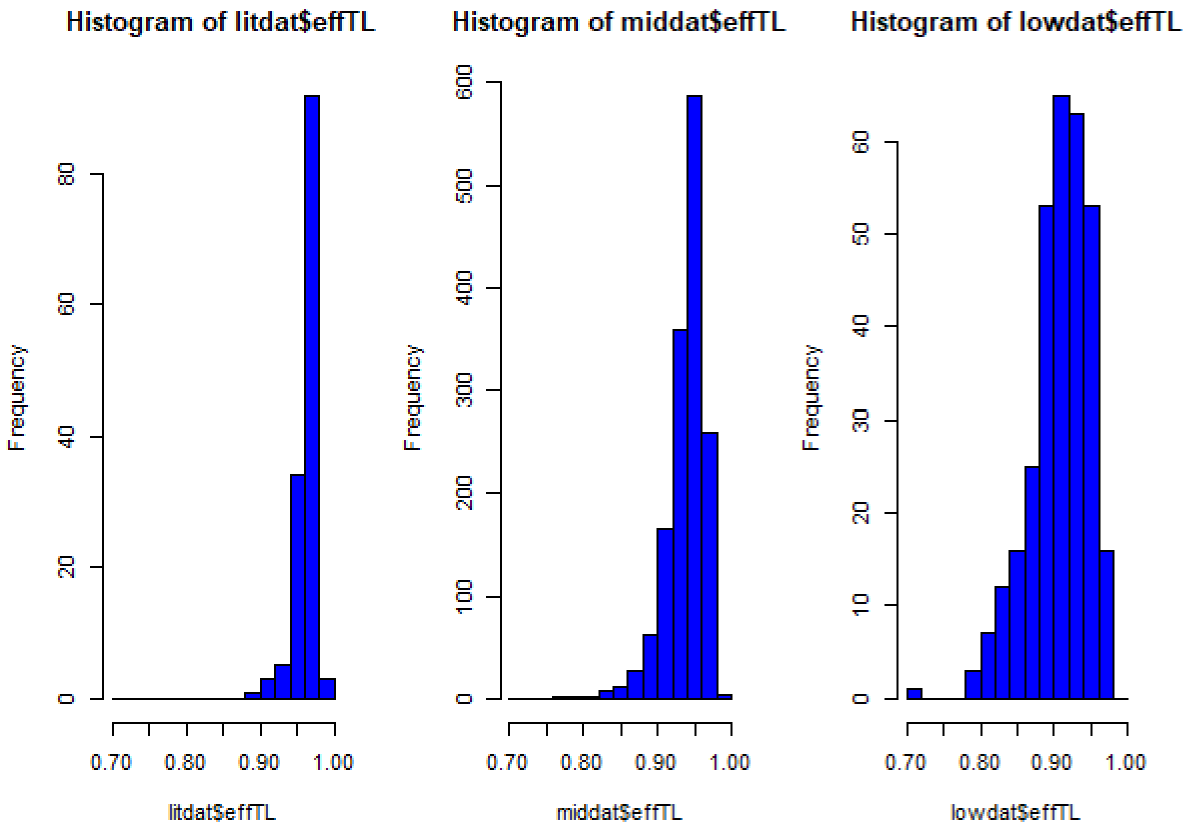
Parlor/AMS coefficient 0,0130

Kiosk coefficient 0,0077

Rail carriage coefficient 0,0042

Sharing by hands coefficient 0,0092

```
> hist(litdat$effTL, breaks=seq(0.70,1.00,0.02), col = "blue",freq = TRUE)
> hist(middat$effTL, breaks=seq(0.70,1.00,0.02), col = "blue",freq = TRUE)
> hist(lowdat$effTL, breaks=seq(0.70,1.00,0.02), col = "blue",freq = TRUE)
```




```
> mean(litdata$effTL)
[1] 0.9612133
> mean(middata$effTL)
[1] 0.9404233
> mean(lowdata$effTL)
[1] 0.9069921
```

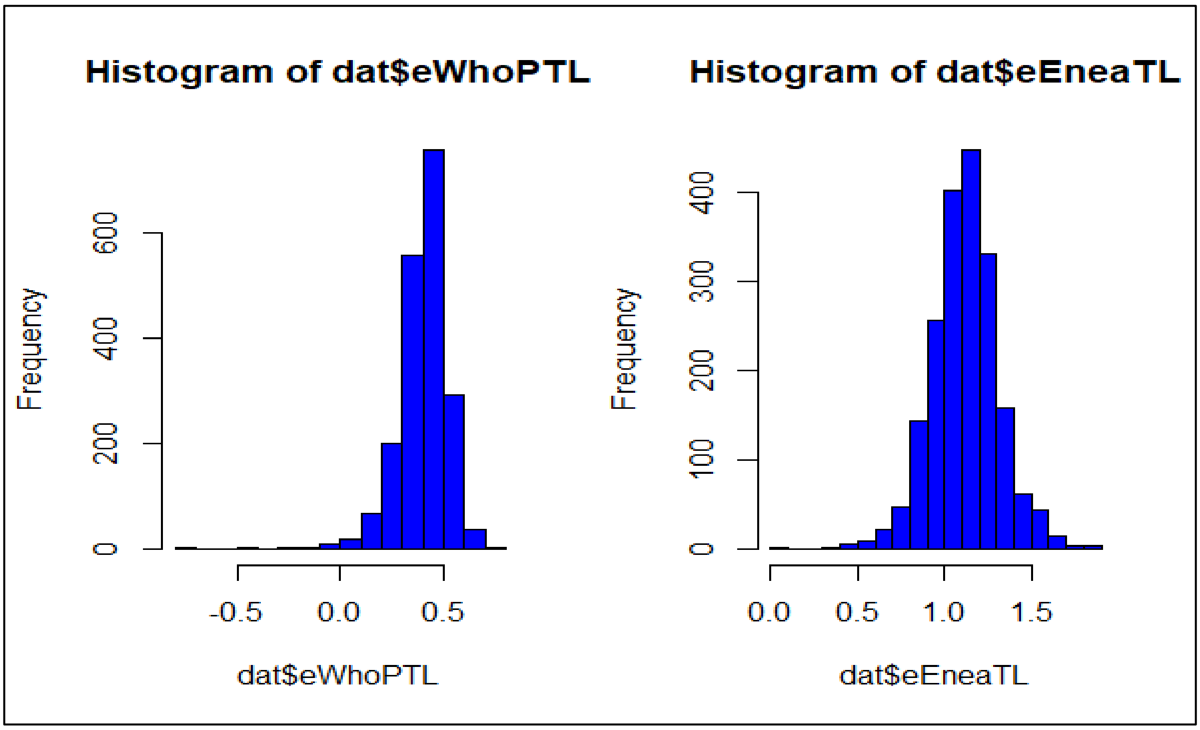
Appendix 4

Energy and protein

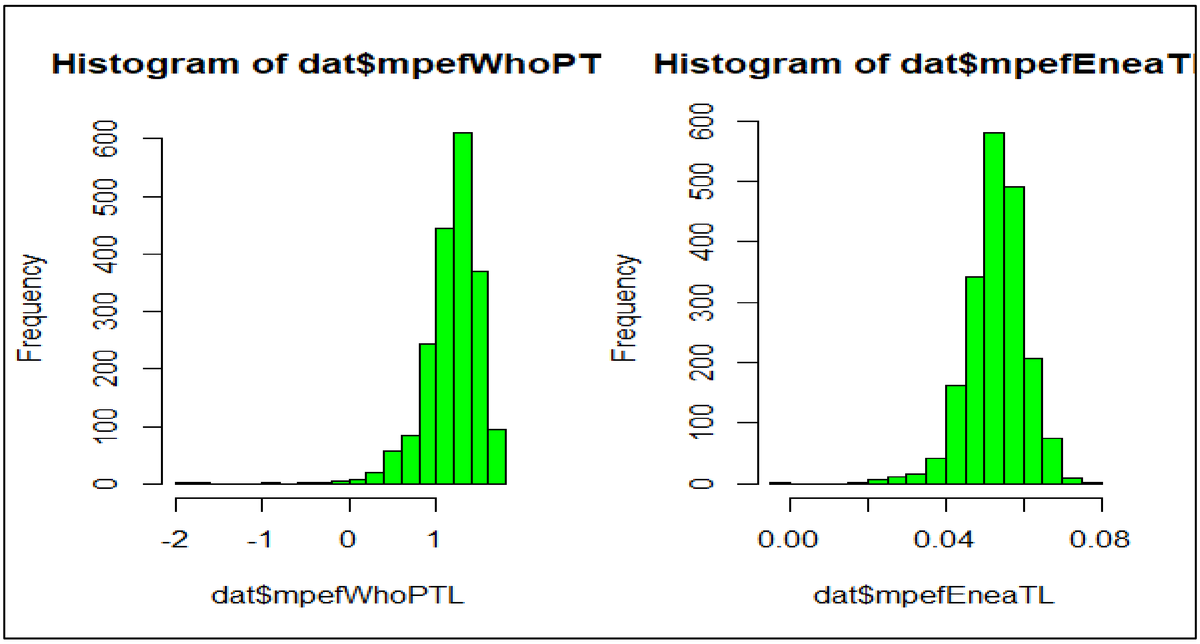
Elasticity of scale at the sample mean

```
> x=0.21161960+0.90833420
> x
[1] 1.119954
```

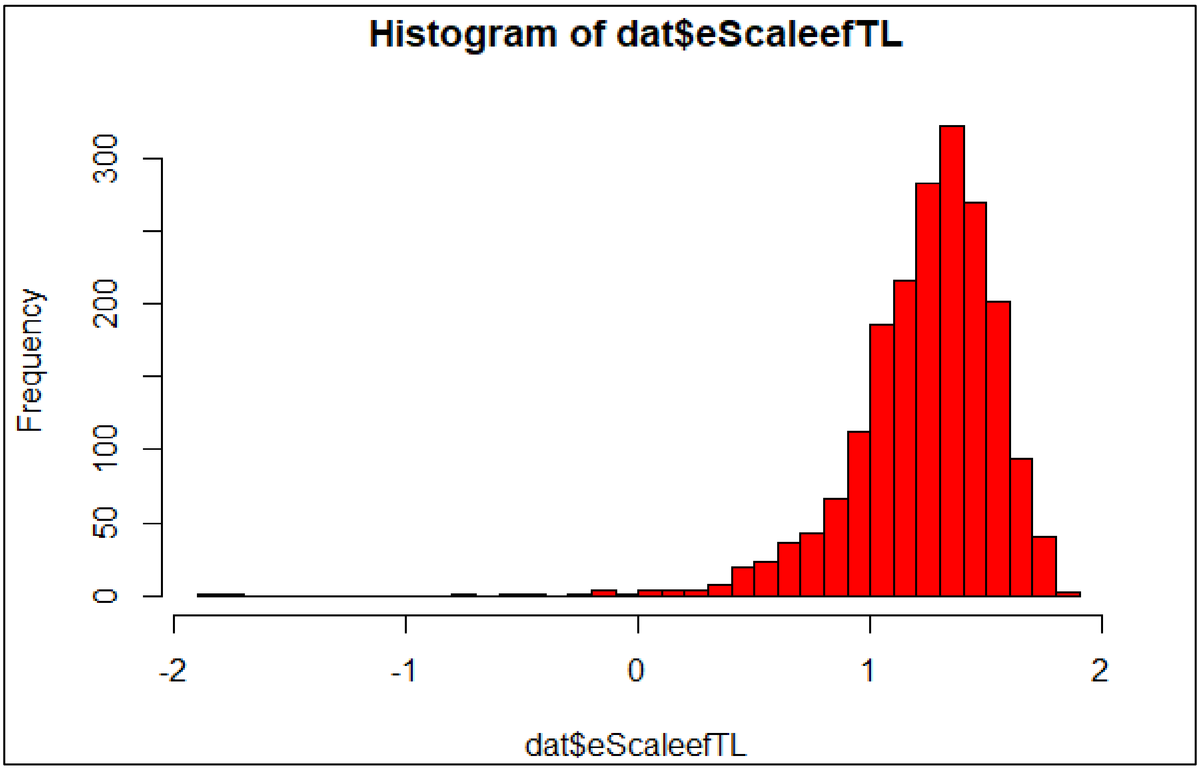
Output elasticities

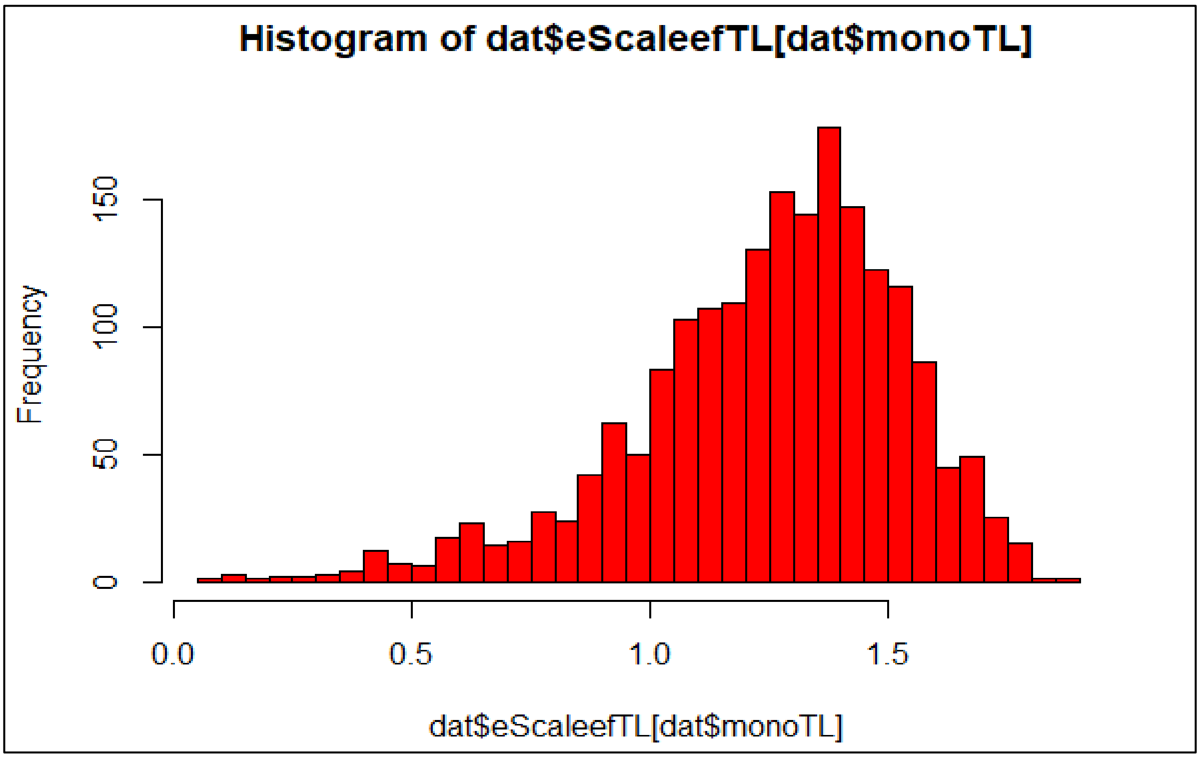


Marginal products



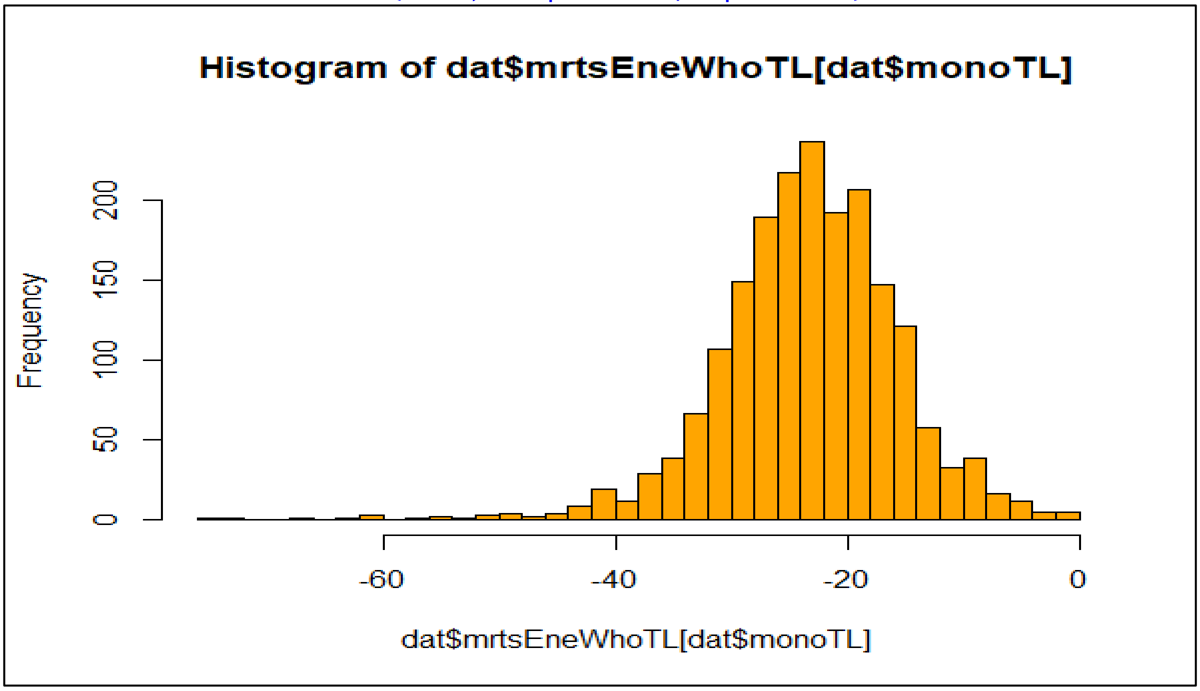
Elasticity of scale





Marginal rates of technical substitution

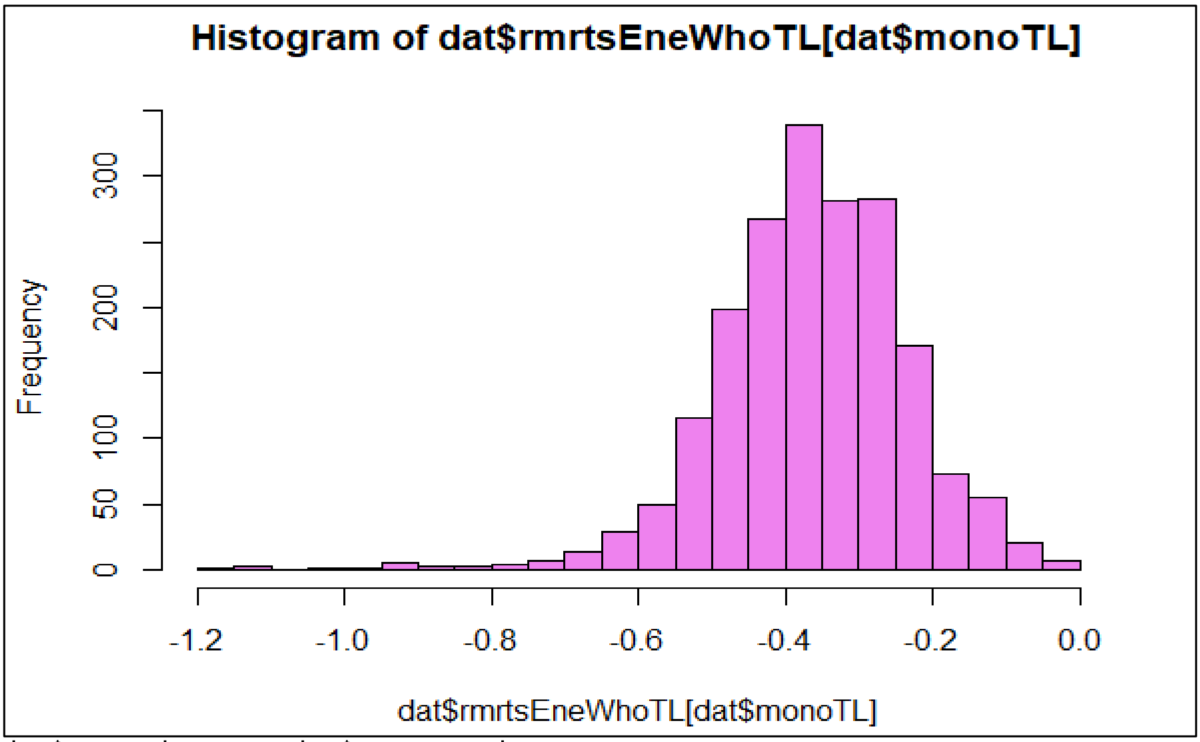
```
dat$mrtsEneWhoTL <- with( dat, - mpWhoPTL / mpEneaTL )
```




```
dat$mrtsWhoEneTL dat$mrtsEneWhoTL
     -0.04352424      -22.97570523
```

Relative marginal rates of technical substitution

```
dat$rmrtsEneWhoTL <- with( dat, - eWhoPTL / eEneaTL)
```




```
dat$rmrtsWhoEneTL dat$rmrtsEneWhoTL
       -2.7832304        -0.3592947
```

Appendix 5

The approximation of prices for energy and protein

Cereal	Protein	Energy	Price
Oat	134	12,3	145
Barley	126	13,2	130
Wheat	125	13,7	142
Pea	230	12,9	210
Rapeseed	358	9,9	257
Fava bean	300	11,9	230
“null”	0	0	0

The contents of protein and energy are from the statistics of Vilja-alan Yhteistyöryhmä and Natural Resources Institute Finland (LUKE). The prices are from Markkinakatsaus of Maaseudun Tulevaisuus (6.4.2020).

```
Call:
lm(formula = y ~ x1 + x2, data = dat)

Residuals:
         1          2          3          4          5          6          7 
 4.900e-03 -1.000e-02  4.996e-05  1.095e-02 -1.393e-03 -4.746e-03  2.344e-04 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.0002344  0.0081468  -0.029  0.97842    
x1           0.5823579  0.0311771  18.679 4.84e-05 ***
x2           0.0050649  0.0007888   6.421  0.00302 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.008193 on 4 degrees of freedom
Multiple R-squared:  0.9939,Adjusted R-squared:  0.9908 
F-statistic: 324.6 on 2 and 4 DF,  p-value: 3.749e-05
```

Coefficient of protein: 0.582358

Coefficient of energy: 0.005065

Rate of prices: MJ of energy/kg of protein: 1/115

The rate of contents of protein and energy for diet

Based on the statistics of this thesis by the subgroup formed of the MoC:s at the level of 0.97 or more efficiency:

```
(mean(hidata$Energyamount))/(mean(hidata$Wholeprotein))
[1] 64.49742
```

Based on the Ruokintasuunnitelma of ProAgria (15.4.2020):

Dieetin koostumus		keskim. Lypsävät 29.79	15	20	25	30	35	40	45
Väkirehua kg/maito kg			0,25	0,31	0,36	0,42	0,40	0,39	0,39
Kuiva-aineen syönti		**21,75**	13,41	16,12	19,04	21,80	24,84	27,68	30,39
Korjattu ME	MJ/kg ka	**10,9**	11,3	11,1	11,0	10,9	10,8	10,8	10,7
Korjaamaton ME	MJ/kg ka	**11,73**	11,30	11,45	11,60	11,73	11,73	11,72	11,74
OIV	g/kg ka	**97**	94	95	96	97	97	97	97
Raakavalkuainen	g/kg ka	**167**	167	167	167	167	167	167	167
Tärkkelys	g/kg ka	**267**	132	180	223	268	265	265	268
Karkearehun kuitu	g/kg ka	**257**	385	340	299	256	259	259	257
Väkirehun osuus	ka:ssa	**0,50**	0,25	0,33	0,42	0,50	0,49	0,49	0,50
Syönti, kg/eläin/pv		**65**	53	58	63	66	75	84	91

```
(10.9/0.167)
[1] 65.26946
```